

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2002, 05:31:50 ; Search time 2313.83 Seconds  
(without alignments)  
16279.401 Million cell updates/sec

Title: US-09-758-269-5  
Perfect score: 1800  
Sequence: 1 atggcttttcacgcaac.....tggcgaagcaggctgtgtga 1800

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues  
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pi.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_ats.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pi.\*  
26: em\_ro.\*  
27: em\_ats.\*  
28: em\_un.\*  
29: em\_vl.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htgo\_inv.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
------------	-------	-------	-------	--------	----	-------------

1	1800	100.0	1800	6	AX148306	AX148306 Sequence
2	1800	100.0	3869	8	AB026549	AB026549 Arabidops
3	1800	100.0	52232	8	AB028617	AB028617 Arabidops
c	1798.4	99.9	2331	8	AY056255	AY056255 Arabidops
5	862.8	47.9	95769	8	AC013430	AC013430 Genomic s
6	824.6	45.8	1818	6	AX148316	AX148316 Sequence
7	824.6	45.8	2171	8	LENEOKANT	Z97215 Lycopersico
8	809.4	45.0	2164	8	STU276244	AJ276244 Solanum t
9	798.4	44.4	1839	6	AX148312	AX148312 Sequence
10	798.4	44.4	2432	8	AB030293	AB030293 Vigna ung
c	791.6	44.0	2398	8	AF190462	AF190462 Phaseolus
12	750.8	41.7	72058	8	AC074176	AC074176 Arabidops
13	738.2	41.0	1752	6	AX148302	AX148302 Sequence
14	738.2	41.0	110102	8	ATF28512	AL021710 Arabidops
15	738.2	41.0	194143	8	ATCHRIV48	AL161548 Arabidops
16	696.8	38.7	2153	8	AF224672	AF224672 Persea am
17	665	36.9	2310	8	AF224671	AF224671 Persea am
18	606.6	33.7	1815	6	AX148314	AX148314 Sequence
19	606.6	33.7	2498	8	ZMU95953	U95953 Zea mays vi
20	528.2	29.3	1734	6	AX148310	AX148310 Sequence
21	528.2	29.3	78921	8	AB028621	AB028621 Arabidops
22	498.4	27.7	82697	8	ATT9A21	AL021713 Arabidops
23	314.6	17.5	602	8	AF399655	AF399655 Lactuca s
24	177.4	9.9	1971	8	WD293765	Z93765 M. domestica
25	169.6	9.4	1788	6	AX148304	AX148304 Sequence
26	169.6	9.4	2064	8	AY056789	AY056789 Arabidops
27	169.6	9.4	96574	8	ATT18B16	AL021687 Arabidops
28	169.6	9.4	199382	8	ATCHRIV50	AL161550 Arabidops
29	169.2	9.4	1950	6	AX085456	AX085456 Sequence
30	155.8	8.7	1769	8	AY029525	AY029525 Phaseolus
c	143.6	8.0	122241	2	AP004129	AP004129 Oryza sat
32	139.8	7.8	1617	6	AX148308	AX148308 Sequence
33	139.8	7.8	1617	6	AX148318	AX148318 Sequence
34	139.8	7.8	1846	8	ATH5813	AJ005813 Arabidops
35	106.8	5.9	1967	8	AF224670	AF224670 Persea am
c	71.4	4.0	630	8	CSA416714	AJ416714 Crocus sa
37	63.2	3.5	138541	2	OSJN00215	AL663011 Oryza sat
38	57.8	3.2	34869	1	AF324838	AF324838 Streptomy
c	57.2	3.2	657	8	CSA416712	AJ416712 Crocus sa
40	56.2	3.1	39428	1	AF322256	AF322256 Streptomy
41	56	3.1	84730	2	AC074054	AC074054 Oryza sat
c	56	3.1	140864	2	AC074355	AC074355 Oryza sat
c	55.6	3.1	15226	1	AE006980	AE006980 Mycobacte
c	55.6	3.1	24680	1	MTY21C12	Z95210 Mycobacteri
45	54.8	3.0	187126	2	AP004729	AP004729 Oryza sat

ALIGNMENTS

RESULT 1  
AX148306  
LOCUS AX148306 1800 bp linear PAT 08-JUN-2001  
DEFINITION Sequence 5 from Patent, EP116794.  
ACCESSION AX148306  
VERSION AX148306.1 GI:14347193  
KEYWORDS thale cress.  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsia.

REFERENCE 1 (bases 1 to 1800)  
AUTHORS Iuchi, S., Kobayashi, M. and Shinozaki, K.  
TITLE Transgenic plants carrying neoxanthin cleavage enzyme gene  
JOURNAL Patent: EP 1116794-A 5 18-JUL-2001;  
Riken (JP)

FEATURES  
Location/Qualifiers  
source 1..1800  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
CDS 1..1800  
/note="unnamed protein product"

```
/codon_start=1
/protin_id="CAC41194.1"
/db_xref="GI:14347194"
/translation="MASFTATAVSGRWLGNHTQPLSSQSSDLSYCSLPMASRV
TRKUNSSALHTPPALHPFKOSSNSPAIVVKPKAKESNTKOMNLFORAAALDAABG
FLVSHKELPLPKTADPSVOIAGNFAPVNEQPVRRNLPVVKGLPDSIKGVYVRNGANP
LHEPVTGHHFPDGDGMVHAKVFEHGSAYACRFQTNRFVQERQLGRPFKPAIGELH
GHTGIAELMLFYAPARAAGIYDPAHGTCTVANAGLVYFNGLLAMSEDDLPHYVOITPNG
DLKTVGRFDDQLESTMIAPHKVDPSGELFALSYDVVSKPLYKIFRFPSPDGTKSPD
VEIOLDQPTWMHDFAITENFVVPDQVQVFLPEMIRGSPVVDKNNKVARFGLLDKY
AEDSSNIKWIDAPCFCHLWNAWEEPEDEVVIGSCMTFPPDSIFNESDENLKSVLS
EIRLNLTKGSTRPIYLSNEDQOVNLEAGMVRNMLGRKTKFAYLALAEWPVKVSGFA
KVDLITCEVKKHLVGNRYGGEPLFLPGEGEDEGVILCFVHDEKTKWSELQIVNAV
SLEVEATVKLPSRVPYGFHGTFFGADDLAKVW"
BASE COUNT      458 a  464 c  439 g  439 t
ORIGIN

Query Match      100.0%; Score 1800; DB 6; Length 1800;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1800; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atggctttttcacggcaacggctgctgtttctggagatggcttggcgcaatcatact 60
Db 1 ATGGCTTCTTTCACGGCAACGGCTCGCTTCTGGGAGATGGCTTGGTGGCAATCATACT 60

Qy 61 cagcggccattatgctcttccaaagctccgacttgatgtattgttagtctcttaccatg 120
Db 61 CAGCGCCATTATGCTCTTCTCAAAGCTCCGACTTGAATTATGTAGTCTCTTACCTATG 120

Qy 121 gccagtgtgtcacacgttaagcttcaatctcgtctgcttccacactcctccagctctt 180
Db 121 GCCAGTGTGTGACACGTAAGCTCAATGTTTCATCTGCGCTTCACACTCTCCAGCTCTT 180

Qy 181 catttccctaagcaatcatcaaaccttcccgcattgtgttaagcccaagccaaagaa 240
Db 181 CATTTCCCTAAGCAATCATCAAACCTCTCCGCCATTGTTGTTAAGCCCAAGCCAAAGAA 240

Qy 241 tccaaactaaacagatgaatttgttcagagagcggcggcggcagcgttggagcggcg 300
Db 241 TCCAACTAAACAGATGAATTTGTTCAGAGAGCGCGCGCGCAGCGCTTGAGCGCGCG 300

Qy 301 gaggttccctgtcagccacgagagactacaccgccttcttcaaacggctgacacctagt 360
Db 301 GAGGTTTCTTGTGAGCCACAGAGAGCTACACCCGCTTCTTAAACCGGCTGATCCTAGT 360

Qy 361 gttcagatcgccggaaattttgtctcgggtgaatgaacagccgctccggcgttaattctccg 420
Db 361 GTTCAGATCGCGGAAATTTGTCTCCGGTGAATGAACAGCCGCTCCGCGGTAAATCTTCCG 420

Qy 421 gtggtcggaaaaacttcccgattccatcaaaagagtgatgtgcgaacggagctaaacca 480
Db 421 GTGGTCGGAAAACTTCCCGATTCCATCAAAAGAGTGTATGTGCGCAACGGAGCTAAACCCA 480

Qy 481 cttcacgagcgggtgacagggtaccactcttcacgagagagcgtatgtttcacgcgctc 540
Db 481 CTTACAGAGCGGTTACAGGTACACACTTCTTCGACGAGAGCGGTATGGTTACGCGGCTC 540

Qy 541 aaatttcgaacacgggttcagctagctacgcttgcgggttactcagactaaacgggtttgt 600
Db 541 AAATTCGAACACGGTTACGCTAGCTACGCTTGC CGGTTTACTCAGACTTAA CCGGTTTGT 600

Qy 601 caggaaactcaattgggtcgaccggttttccccaagccatcggtagcttcacggccac 660
Db 601 CAGGAACGTCAATTTGGGTGACGCGGTTTTTCCCAAAGCCATCGGTGAGCTTTCACGGCCAC 660

Qy 661 accgatttgcggaactcatgctattcacgccaagcgtgagcgggtatagtcgaccgcg 720
Db 661 ACCGTAATGCGCGACTCATGCTATCTTACGCCAGAGCTGACGCGGTTATAGTCACCCCG 720

Qy 721 gcacacggaacgggtgtagctaaacgggttttggcttatttcaatggccgggtattggct 780
Db 721 GCACACGGAACGGGTGAGCTAAACGGGTTTGGTCTATTTCATGCGCGGTTATTGGCT 780
```

```
Qy 781 atgtcggagagatgatttaccttaccgaattccagatcaactcccaatcgagatttaaaacc 840
Db 781 ATGTCCGAGAGATGATTTACCTTTACCAAGTTCAGATCACTCCCAATGAGATTTAAAAACC 840

Qy 841 gttggtcgggttcgattttgatggacaattagaatcccaatgattgccaccgccgaagtc 900
Db 841 GTTGGTTCGGTTCGATTTTGATGGACAATTAGAATCCACAATGATTGGCCACCAGAAATGC 900

Qy 901 gaccggaaatccgggtgaactcttgcgttttaagctacgactcgctttcaagccttacct 960
Db 901 GACCCGGAATCCGGTGAACCTCTTCGCTTTTAAGCTACGACGCTCGTTTCAAAAGCCTTACCTA 960

Qy 961 aaatacttcgattctcacggagcgaactaaatacaccggacgtcgagattcaactgat 1020
Db 961 AAATACTTCCGATTCTCACCGACGGAACTAAATCACCGACGTCGAGATTTCACTTGAT 1020

Qy 1021 cagccaaacgatgatgcacgatcttcgcgattacagagaacttcgtcgtacactgaccag 1080
Db 1021 CAGCAACGATGATGCACGATTTCCGATTTACAGAGAACTTCGTCGCTGCTACCTGACCCAG 1080

Qy 1081 caagtctgtttcaagctgccggagatgatccgggtgggttccgggtggttttacgacaag 1140
Db 1081 CAAGTCGTTTTCAAGCTGCCGAGATGATCCGCGTGGGTCTCCGGTGGTTTTACGACAAG 1140

Qy 1141 aacaagtcgcaagattcgggatttttagaacaatacgcgaagatttcatcgaacattaa 1200
Db 1141 AACAAAGTCCGAAGATTCGGGATTTTAGACAAATACCCGCAAGATTCATCGAAATTAAAG 1200

Qy 1201 tggattgatgtccagattgtcttctgcttccatctctggaaacgcttgggaagagccagaa 1260
Db 1201 TGGATTGATGCTCCAGATTGCTTCTGCTTCCATCTCTGGAACGCTTGGGAAGAGCCAGAA 1260

Qy 1261 acagatgaagtctcgtgatagggtcctgtatgactccacagactcaattttcaacgag 1320
Db 1261 ACAGATGAAGTCGTCGTGATAGGTCCTGTATGATCTCCACAGACTCAATTTTCAACGAG 1320

Qy 1321 tctgacagagaatctcaagagtgctctgtctgaaactcgcctgaactcctcaaacgggtgaa 1380
Db 1321 TCTACAGAGAATCTCAAGAGTGTCTGTCTGAAATCCGCTGAAATCTCAAAACCCGGTAA 1380

Qy 1381 tcaactcgcgcgtccgatcatctccaacgaagatcaacaagtcaacctcgaagcaggatg 1440
Db 1381 TCAACTCGCGTCGATCATCTCCACGAGATCAACAAGTCAACCTCGAAGCAGGGATG 1440

Qy 1441 gtcaacgaaacatgctcggcgttaaacccaaatctgcttacttcttagcgcgagccg 1500
Db 1441 GTCAACGAAACATGCTCGCGCGTAAACCAAATTCGCTTACTTGGCTTTAGCCGAGCCG 1500

Qy 1501 tggcctaaagtctcaggatttcgctaaagttagctcactactggagaagtaagaacat 1560
Db 1501 TGGCCTAAAGTCTCAGGATTTCGCTTAAAGTTGATCTCACTACTGGAGAACTTAAAGAACAT 1560

Qy 1561 ctttacggcgataaccgttacggagggagagcctctgtctcccgagagagggagagag 1620
Db 1561 CTTTACGGCGATAACCGTTACGGAGGAGAGCCCTCTGTCTTCTCCCGGAGAGAGGAGAG 1620

Qy 1621 gaagacgaaggaatacatcctctgttttcggttcacgacgagaagacatggaatcgagtt 1680
Db 1621 GAAGACGAAGGATACATCTCTGTTCGTTCAACGACGAGAAGACATGGAATTCGAGTTA 1680

Qy 1681 cagatagttaacgcggttagcttagaggttgaagcaacgggttaaaacttcgtaagggtt 1740
Db 1681 CAGATAGTTAACGCGGTTAGCTTAGAGGTTGAAGCAACGGTTAAACTTCGCTCAAGGGTT 1740

Qy 1741 ccgtacggatttcacggtacattcatcgagcggatgatttggcgaagcaggtcggtgga 1800
Db 1741 CCGTACGGATTTACGGGTACATTCATCGAGCCGATGATTGTGGCGAAGCAGGTCTGTGTA 1800

RESULT 2
AB026549 3869 bp DNA linear PLN 24-OCT-2001
LOCUS
```

```

DEFINITION Arabidopsis thaliana gene for neoxanthin cleavage enzyme, complete
            cds.
ACCESSION AB026549
VERSION   AB026549.1 GI:16416373
KEYWORDS  neoxanthin cleavage enzyme.
SOURCE    Arabidopsis thaliana (strain:Col.) DNA.
ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 3869)
AUTHORS   Iuchi,S., Kobayashi,M. and Shinozaki,K.
TITLE     Characterization of neoxanthin cleavage enzyme from Arabidopsis
            thaliana
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 3869)
AUTHORS   Iuchi,S. and Shinozaki,K.
TITLE     Direct Submission
JOURNAL   Submitted (22-APR-1999) Satoshi Iuchi, RIKEN, Plant Mol. Bio.;
            3-1-1 Koyuadai, Tsukuba 305-0074, Japan
            (E-mail:luchi@tc.riken.go.jp, Tel:81-298-36-4359)
FEATURES   Location/Qualifiers
            1..3869
             /organism="Arabidopsis thaliana"
             /strain="Col."
             /db_xref="taxon:3702"
             1831..3830
              /codon_start=1
              /product="neoxanthin cleavage enzyme"
              /protein_id="BAB70609.1"
              /db_xref="GI:16416374"
              /translation="MASFTATAASQVRLGNGHTQPLSSQSSDLSSYSSSLPMASRV
                TRLVNSALLHTPALHPKQSSPAIVVVKPAKSNTKOMLRFORAAAALDAAGP
                FLVSHKLLPLTADPSVOIAGNPVNEQPVRLPVVGLKLPDSTKGVVVRNGANP
                LHPVTVGHFPDGDGMVHAKVPHGSASYACRFQTNRFVQERQLGRPVPPKALGELH
                GHTGARLMFLYARAAGAVDVPAGHTGVAGNAGLVYNGRLUWSEDDLPVQVITPND
                DLKTQGRFDGQGLSELTMIADHPVDPSGSEFALSVDVWSKPLYKFKRSPDGTGKSP
                VETLQDQMTWDFATFENLVNWPVDQVFKLPEMIRGSPSVYVDKNKVARFGILDKY
                AEDSSNKWIDAPDTCFHLUNAWBERPETDEVVVI GSCMTPPDSIFNESDENLKSVL
                EIRLNKLTGSTRPILISNEDOOVNLEAGMVNENMLGRKTKFAYLALAPMPKVSQFA
                VLDTTGVCKHLYGDNRYGGEPLFLPBGEGGEDEGYILCFVHDEKTKWSELOIVNVA
                SLEVAETKVLPSRPVPGHTFEGADLAKQVV"
BASE COUNT 1155 a 845 c 773 g 1096 t
ORIGIN
Query Match 100.0%; Score 1800; DB 8; Length 3869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1800; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggctctcttcaggcaacggtcgggttctctggagatggcttgatggcaatacact 60
    |||||
Db 1831 ATGGCTTCTTTCAGCGCAACGGCTCGCGTTTCTGGGAGATGGCTTGGTGGCAATCACTACT 1890
    |||||

QY 61 caggcgccattatcgctcttctcaaagctccgacttgattttagctctcttaccctatg 120
    |||||
Db 1891 CAGCGCCATTATCGTCTTCTCAAAGCTCCGACTTGAGTTATTGTAGCTCTTACCTATG 1950
    |||||

QY 121 gccagtctgttcacacgaagctcaattgtttcatctcgctttcacactctccagctctt 180
    |||||
Db 1951 GCCAGTCGTGTGCACGCTAGCTCAATGTTTCAATCGCTTTCACGCTTTCACACTCTCTCCAGCTCTT 2010
    |||||

QY 181 catttccctaagcaatcatcaactctccgcattgttcttgaagcccaagccaaagaa 240
    |||||
Db 2011 CATTTCCCTAAGCAATCATCAACTCTCCCGCCATTGTTGTTAAGCCCAAGCCAAAGAA 2070
    |||||

QY 241 tccaactaactaacagatgaattttttccagagagcggcgccgagcgttggagcgcgcg 300
    |||||
Db 2071 TCCAACACTAAACAGATGAATTTGTTCCAGAGACGCGCGCGGAGCGTTGGACGCGCGG 2130
    |||||

QY 301 gagggttctcttgccaccagagaagctacacccgcttctctaaacggctgatcctagt 360
    |||||
Db 2131 GAGGGTTCTCTTGTCACCGACGAGAAGCTACACCGCTTCTCTAAAACGGGTGATCTTAGT 2190
    |||||

```

Qy	1441	gtcaacagaaacatgctcgccgtaaaacaaatcgccttacttgcttgcgcgagccg	1500
Db	3271	gtcaacagaaacatgctcgccgtaaaacaaatcgccttacttgcttgcgcgagccg	3330
Qy	1501	tggcctaagtcctcaggattcgctaaagtgtgatctcactactgagaaagttaagaaacat	1560
Db	3331	tggcctaagtcctcaggattcgctaaagtgtgatctcactactgagaaagttaagaaacat	3390
Qy	1561	ctttacgagcgaataaccgttacgaggaagagcctctgtttctcccggaagaaggagagag	1620
Db	3391	ctttacgagcgaataaccgttacgaggaagagcctctgtttctcccggaagaaggagagag	3450
Qy	1621	gaagcgaagatacatcctctgtttcgttcacgacgagaagacatgaaatcggagatta	1680
Db	3451	gaagcgaagatacatcctctgtttcgttcacgacgagaagacatgaaatcggagatta	3510
Qy	1681	cagatagttaacgcgcgttagcttagaggttgaagcaacggttaaacctccgtcgaagggtt	1740
Db	3511	cagatagttaacgcgcgttagcttagaggttgaagcaacggttataacttccgtcgaagggtt	3570
Qy	1741	cgtacgagatttcacggttacattcatcgcgacgacgagcagattggtggaagcaggtcgtgtga	1800
Db	3571	ccgtacgagatttcacggttacattcatcgcgacgacgagcagattggtggaagcaggtcgtgtga	3630
RESULT 3			
AB028617/C			
LOCUS			
AB028617 Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone:MOA2.			
ACCESSION			
AB028617.1 GI:5041970			
KEYWORDS			
SOURCE			
Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1 clone:MOA2.			
ORGANISM			
Arabidopsis thaliana			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (sites)			
REFERENCE			
AUTHORS			
TITLE			
Sato, S., Nakamura, Y., Kaneko, T., Katoh, T., Asamizu, E. and Tabata, S. Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence features of the regions of 4,504,864 bp covered by sixty P1 and TAC clones			
DNA Res. 7 (2), 131-135 (2000)			
20277480			
2 (bases 1 to 52232)			
Sato, S., Nakamura, Y., Kaneko, T., Kato, T., Asamizu, E. and Tabata, S. Direct Submission			
Submitted (09-JUN-1999) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana Kisarazu, Chiba 292-0812, Japan [E-mail:ynakamu@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934]			
Address for correspondence: kaos@kazusa.or.jp			
For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MOA2			
Genes with similarity to proteins in the databases are described in 'product or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'.			
The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/Grail-1.3/), GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SplicePredictor (Volker Brendel, Stanford University, http://gremlini.zool.iastate.edu/cgi-bin/sp.cgi).			
Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/).			
This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is MLN21 and the 3' clone is MIE1.			
Location/Qualifiers			

source

```
1. .52232
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
/clone="MOA2"
/clone_lib="Mitsui P1"
join(1996. .2132,2225. .2349,2455. .2501,2648. .2725,
2807. .2931,3021. .3087,3216. .3278,3370. .3439,3529. .3629,
3718. .3822,3906. .4091)
/notes="gene_id:MOA2.2"
/codon_start=1
/evidence=not experimental
/product="glycolate oxidase"
/protein_id="BAB01334.1"
/db_xref="GI:11994212"
/translation="MEITNVTEYDAIAKOKLPMVYDYVYASGAEDQWTQENRNFAFAR
IFRILIDVSDIMTITVLGFKISMPIMVAPTAMQMAHPDGEYATARAAGATII
MTLSSWATSEVASTGPIRFOLYVYKRNADVQLVRRAGAFKALATVDTPTP
LGRRESDIKNRFTLPNLTAKNFEGLDLGKMDANDGLASVAGOIDRTLSKDWQM
LQITIKPLIVKGLTGEDARTAIQAGAAGIIVSNHGARQLDYVPATISALEVVKAT
QGRIPVLDGVRGTDVFKALGASGIFGRPVVFSLAAGEAGVRKVLQMLRDEF
ELTMAUSGCRSLKEISRNIHTTEWDTTPRSARL"
complement(join(4591. .4659,4763. .4924))
/notes="gene_id:MOA2.3
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01335.1"
/db_xref="GI:11994213"
/translation="MIGRRAGTNRMRRRDSLLTRFVDSVFYFFRLAEPEILFVLFM
LITVIFKDLTSRPEYNRIIVKEPKGSDIWP"
complement(11870. .13669)
/notes="gene_id:MOA2.4"
/codon_start=1
/evidence=not experimental
/product="9-cis-epoxycarotenoid dioxygenase"
/protein_id="BAB01336.1"
/db_xref="GI:11994214"
/translation="MASPTATAAVSGRWLGNTQPLSSQSDLSYCSSLPMSARV
TRKLNSSALHTPPALFPKQSSNPAIVVKPKAKESNTKQMLFORAAAAADAAAGP
FLVSHKRLHPLPTADPSQIAGNAPVNEQVRRNLPVVGLPDSIKGYVVRGANGP
LHEPVTGHHFFDGDGMVHAVKFEHGSASACRFTQTNRFVQERQLGRPVPKAIGELH
GHTGIARMLFARAAGIVDPAHGTGVANAGLVYFNGLLMSDDLPRYOVOITPNG
DLKTVCRFPDGLQESTMTAHKPVDPSEGELEFSLSDVYKVKYFRSPDGTSPD
VEIQDQPTMMHDFALTENPVVVDQVVKFLPEMIRGSGSPVYDKNKVARFGLDKY
AEDSSNKWIDAPCFCHLNAWEEPTDVIWVIGSMTTPPDSINENSDNSLVLS
EIRLNLTGESTRRPIISNEDQQVNEAGMNRNMLGRKTKFAYLALAEPPKVSQFA
KVDLTTEGVKKHLYGDNRYGGEPLFLPGEGBEDEGYILCFVHDEKTKWSELQIVNAV
SLEVEATVKLPSRPVYGFHGTFRIGADDLAKQVV"
join(29980. .30324,30421. .30528,30616. .30660,30841. .30913,
31021. .31133,31239. .31298,31376. .31492,31588. .31649,
31733. .31793)
/notes="contains similarity to RNA-binding protein
gene_id:MOA2.5"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01337.1"
/db_xref="GI:11994215"
/translation="MATVTPEMTDVVVVDDVKEVSTKSEKIIDEGIEKSSITDKTET
ESRLDMHKLVMFKLNPALKEFPFSPYDPKKNVAKANOQFPADDFETTKQSGBE
FLDLDKDNONTKRRNYSOGRRLTGRIKSAOERDSIRRTVYVSDIDQSVTEGLACL
FNSGQVDCRICDGRSHVRFAPVFPADQGAHEALS LGGTMLGYPVAVLSKTAI
LPNVGTFPLRSEDEREMCTIYCTNIDKKVSQADVRNPFESACGEVIRURLLDQJH
STRIAFVEFALDALSALNSGMVVGSPQIRVSPSKTPVRPRITPPSTN"
complement(32182. .36456)
/notes="gene_id:MOA2.6"
/codon_start=1
/evidence=not experimental
/product="disease resistance complex protein"
/protein_id="BAB01338.1"
/db_xref="GI:11994216"
/translation="MANSYLSSCANVMVVRINTSQBELVELCKGSSALLKRLKVALV
```



TANPLADADORAEBHREVKHWLTGIDKAPFAEDDILDELQTEALRRVVAEAGLGG  
LFQNLMAQRIQKKEIPMEKVRLLLEHHVHKHLEVIKGEYSSETPQWRQASRRP  
DDLQGRUGVREVDKLNLLDSEDEISIGKPAVIVSVNGPVGKTKLTLEIVENDYR  
VHEFVEMQVTSAGINFNVTFTVAQVQDITSSANVEDLPSLQIQKLTQKLMNEBC  
VLDDPDESDESEFVOAFVDAEESKIVLTTTSEIVSTVAEAKIYQMKLMTNEC  
WELISRFAGNISVGSINOEGEGIKRIAEQCKGLPLAARAIASHLSKKNPDWYAV  
KNLSSYTNLSILPVKLSYDLSPLQKRCFCALCIFPKGHVDFDEELVLLMADLLY  
QPSRRRLDIGNDYLGDVAQSFQRDITITPMVMDLMDLAKAVSGDFCFLRED  
DNIPEITSTRHSYSQCDASVAFRSICGAELFRLTLPFNPSPTSLESQLEKVLN  
PLNALSGDRLTSLSHYQITNLKSLKGLKLLRYLDLSSTKIKELPBFVCTLCNLQTL  
LLSCRODLSKPIAGIELINLNDLVGTPLVEMPPGIIKLSRLQKLSLNGVGRSGA  
GLHELKSHLRGLRLSELONVAFASEAKDAGIKRPFLLDGLIKLWTVKSGFVPS  
FNALACQKEVLRMLEBPHLKTICISYQCGAPFKWLGDSFFGITSVLSSNLCI  
SLPVGQPLKYLUSIKFENILQKVGLDFFPGENNSGVPPQSLQILKPYGMPRWDW  
CITPELEGIPFCLQKLIQRCPSLRKFPPEGLSPSTEVITSDCLRAVSGGENFRRS  
LNTNPESPASIPMSRRELSSTGNPKSDASTAQPGFASSSQNDNVEVTSLSLS  
LPKRDQTFDPOYTQSLGSLPOQFEPAVISAARYSGVLSIDIPSTLSPMSRTSLVPP  
KNEGSLIPGSSSYOHYOGIKSSVSPSRSEAIKPSQYDDDETMEVLYKTDISHME  
LPQNLKSHIDSQGLSLPENLSTESPNLHELILIACHLESFPGHPPTLKLXI  
RDCKUMFTESLQSRYSQLEYLFISSCSNVLNPLSLFPFKLSRSIIRDCESFKTF  
SIHAGLDGDRITALESLIRDCPNLETPQGLPTPKLSMLLSNCKKLOALPEKIFGL  
TSLSLPTIKCPEITIPGGFPNLRITLCISLQKLTPTIEMGLRDLNLRNLEIDG  
GNEDESPFERGLPKSVFSLIRSFENLKTNLKRGFDTKAIEITMELSGCKLQISL  
DEDLPLSCLRISSCSLLTETFAEVETEFPFKVLNIPYVEIDGEIFS"  
38132..41236

CDS

/note="gene\_id:MOA2.7"  
/codon\_start=1  
/evidence=not\_experimental  
/product="disease resistance  
/protein\_id="BAB01339.1"  
/db\_xref="GI:11994217"

/translation="MTGIGEMFLAAFLQALFQTLVSPFRFFFKRBLNENLRLST"  
SNLRQLGRMSLGDFTNGSEHLETRLEKVTIRLERLASORNILGKELTAMIPKOR  
LPTTSLVDESVFVGRDDDDKDIRMFLIPENKONGITVVAIVGIGVGKTKTLLSYLN  
DOHVSQKLGPTWMAHVSEEDPVFKITKVYESVTRCEPTDLDVLVYKLERLTQPL  
LPFLVLLDNLWENFADWLLRQFFIIHAAQSQILVTRISQVASIMCAVHVHNLQPL  
SDGWSLFMKTVFNQEPNLEIGDLAERIHKRCGLPLAVTKLGVLRFEKGVLE  
WERSLRINDLPADKNLNLPLRVSYIYLPALHUKRCFAYCIPFKGHAPEKDKVILL  
WMAEGFQQLTQSSNKLBNELGNEYSELESRLLOKTKRTYIMHDFINELAQFASGES  
SKFEDGCKQVRSRTRVLYLRDNYABEMPEALREVKFLRTLPLSLTWSRSCLLD  
OMVSEKLLPTLRVLVLSLHYKARLPDPFKNISARFLDLSRLEKLPKSLCYM  
YNLQTLVLSYSSLUKELPTDISINLIRYLDLIGTKURQMPRRRGRUKSLOTLTFV  
SASDGRSISELGHLDHGLKIVELQVRVDVADAANLNSKKHLREIDFVWRTGSS  
SENNTNPTQNEAEVFEKLRPHRIEKLAIERKGRFPDMLSDPSFRSIVCIRLR  
ECQYCTSLPSLQGLPKELHISGMVGLQSIGRFYFSDOQLRDQDQOPRSLETLRF  
DNLPMOEWLNDVTRGDLPESLAKLPIRCPELTGTLPTFELSLSLHLYKGLLDF  
QPDHYSYRNQLTSLISKSCDITLVKPLNHPANLKDLEVDQTSLSLSLSLNLHRLG  
PNAURNKIDCQNLQLPLKUNALPQNLQVLTITNCRYLQPMEQQPYHHHPHLPRS  
NVGSPKSHGSHRSYDSSRSRYD"  
42345..42872  
/note="gene\_id:MOA2.8  
/unknown\_protein"

CDS

/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="BAB01340.1"  
/db\_xref="GI:11994218"  
/translation="MSDLVIALVAFVLFVVFIILCIMGKXDVHSPPLPRPQLRL  
MPVYTKDQDGGGTALISVAGTDIYGGSGVGGHGGRIITSDTKLPSPPPPSLP  
PRPVQRPKSVSRDKRELIDTGTAMLVASSLLTSGSASSCGSGSHGGCGGGGGGG  
GGUGGGCGGGGGCGG"  
complement(join(42980..43465,43561..43668,43952..44119,  
44215..44460,44539..44677,44759..44977,45060..45754))  
/note="gene\_id:MOA2.9"  
/codon\_start=1  
/evidence=not\_experimental  
/product="terpene synthase-like protein"  
/protein\_id="BAB01341.1"  
/db\_xref="GI:11994219"

CDS

/translation="MKPKVRDMLKSSKNSDNERIRLJHLMLNLGIAVHFIEIDELIG  
QAFGNLDDITAKENDLETISTMPEVFLRGYMPCKSIGDLSHICIHLYNLAFDFT  
VDANRFKDGRELAEDIRGMQLQYEAHLGTPSEIDMDLAESTFYRLESLS  
NHTATSPHLSKHIQNALYRARVHNLEILVAREYISFYQEEDHDEHTLLKFAKLNFY

COLHYIOELKDLTKWKKBELASKLPIYRIRIVEYFGALALYFEPYSLGRIIVTKI  
TMIVTVDNDAYGTLPEISLVDSQRWDLGDIEKLPSYVKTVFPGVFTLBEIEQ  
EMRPOQRSRYVOAVDAIEIKLGLAYLAISKWARASHVPTFEYEFNGQTSMDHFAAY  
SFTAMEDCDENOTCEWYKSRPKMEALNGVFRINDINTPEQMSRGEVAKGLANCYMK  
OHGVSKEEATGEMNKIYSNVYKIMEEYLTATVPRDILETFDMLNSKPLIJKTLTLM  
FRLVNDVGTTEVHLKLTETISREEVANGLDFYMKOHGVTKKEASQEURKMMKNYKV  
MEBFMNTDHLPLHQVFLURCHNIARIFDVFTEVDGYCDPKGIENFTMSLYLHPINTP  
FIIMTWNTHQYSRISIRGCCIEIKKKKKVVLCSLLIHRYSNYSLVHVLRCFVFLYNSIFLC

Query Match 100.0%; Score 1800; DB 8; Length 52232;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1800; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	atgcttcttttcacggcaacggctgctgtcttcctcgaagctcgacttgacttatttagctcttcaactatg	60
Dy	13669	ATGGCTCTTTTTCACGGCAACGGCTGCGGTTTCTGGAGATGGCTTGGTGGCAATCATACT	13610
Qy	61	cagcgcacattatcgctctctcctcgaagctcgacttgacttatttagctcttcaactatg	120
Dy	13609	CAGCGGCCATTATCGTCTTCTCAAGCTCCGACTTGAGTTATGTAGTCTCTTACCTATG	13550
Qy	121	gccagtctgtcacacgttaagctcaatgtttcctcctcgcttcacactcctccagctctt	180
Dy	13549	GCCAGTCTGTGTACACGTAAAGCTCAATGTTTCACTGCGCTTCACACTCCTCCAGCTCTT	13490
Qy	181	catttccctaaagcaatcatcaactctcccgccattctgtttaagcccaagcccaagaa	240
Dy	13489	CATTTCCCTAAGCAATCATCAAACTCTCCCGCCATTGTTGTTAAGCCCAAGCAAGAA	13430
Qy	241	tccaaactaaacagatgaattgttccagagagcggcggcggcagcgttggagcggcgg	300
Dy	13429	TCCAACACTAAACAGATGAATTTGTTCCAGAGAGCGGCGGCGAGCGCTTGGACGCGCG	13370
Qy	301	gagggtttctctgtcagccacgagaagctacacccgcttctctaaaagcgtgactcctagt	360
Dy	13369	GAGGTTTCTTGTGAGCCACGAGAAGCTACCCCGCTTCTTAAACAGCGCTGATCCTAGT	13310
Qy	361	gttcagatcgccggaattttgtctccggtgaatgaacagcgcctccggcgttaattcttcg	420
Dy	13309	GTTTCAGATCGCGGAAATTTTGTCTCCGCTGAATGAACAGCCGCTCCGCGGTAACTTCCG	13250
Qy	421	gtgctcggaataactcccgattccatcaaaagagtgatgtgcgcaacggagcgttaaccca	480
Dy	13249	GTGCTCGGAAACTTCCCGATTCCATCAAGAGAGTGTATGTGCGCAACGGAGCTAACCCA	13190
Qy	481	cttcacagccggtgacaggtcacacttcttcagcggagcggataggttcagcgcctc	540
Dy	13189	CTTCACGAGCCGGTGACAGGTCAACACTTCTTCGACGAGACGGTATGGTTACGCCCGTC	13130
Qy	541	aaattcgaacacgggttcagctagctacgcttgcgggtttactcagactaacgggtttgtt	600
Dy	13129	AAATTGGAACACGGTTACGCTAGCTACGTTGCGGTTTACTCAGACTAACCGGTTTGT	13070
Qy	601	caggaacgtcaattgggtcgaccgggttttcccaaaagccatcggtgagcttcacggccac	660
Dy	13069	CAGGAACGTCAATTGGGTGCGACGGTTTTTCCCAAAGCCATCGGTGAGCTTTCACGGCCAC	13010
Qy	661	accggtattgcccactcatgctattctacgcagagctgcagcggatagtcgacccg	720
Dy	13009	ACCGGTATTGCCCCGACTCATGCTATTCTACGCCAGAGCTGACGCCGGTATAGTCACCCCG	12950
Qy	721	gcacacggaacgggtgtagctaaacggcgggtttggtctatttcaatggcgggttattggct	780
Dy	12949	GCACACGGAACGGGTGAGCTAAACGGCGTTTGGTCTATTTCATGGCGGTTATTGGCT	12890
Qy	781	atgtcggaggatgatttacccttaccaggttcagatcactcccaatggagatttaaaacc	840
Dy	12889	ATGTGCGGAGGATGATTTACCTTACCAAGTTCAGATCACTCCCAATGGAGATTTAAAAACC	12830
Qy	841	gttggtcgggttcgattttagtgagcaattagaatcccaaatgattgccccaccgaaagtc	900
Dy	12829	GTTTGGTCCGTTTCGATTTTGTATGGACAATTAGAATCCCAATGATTTGCCACCCGAAAGTC	12770

```
Qy 901 gacccggaatccggtgaactcttcgcttaagctacgacgctgcttcaaaccttaccta 960
Db 12769 GACCCGGAATCGGTGAATCTCTCGCTTTAAGCTACGACGTCGTTTCAAGCCTTACCTA 12710
Qy 961 aaatacttccgattctcaccggacggaactaaataaccggagctcgagattcagcttgat 1020
Db 12709 AAATACTTCCGATTCTCACCGACGGAATAAATCACCGGACGTCGAGATTTCAGCTTGAT 12650
Qy 1021 cagccaacgatgacgacgatttcgctgattacagagaacttcgctcgctgacctgaccag 1080
Db 12649 CAGCCAACGATGATGCACGATTTCGCGATTACAGAGAATCTCGTCGTCGTCACCGACG 12590
Qy 1081 caagtcgcttttcaagtcgagagatgacccgctggttcggttcggttcggttcgacacaag 1140
Db 12589 CAAGTCGTTTTCAAGCTCCCGGAGATGATCCCGGTGGGTCTCCGGTGGTTTACGACAA 12530
Qy 1141 aacaaggttcgaagattcgggatttttagacaataatcgcggaagatttcatacgaattaa 1200
Db 12529 AACAAAGTTCGGAAGATTCCGGATTTTAGACAATAACGCCGAAGATTTCATCGAATTAAG 12470
Qy 1201 tggattgatgtccagattgcttctgcttccatctcttggaaacgcttgggaagaccagaa 1260
Db 12469 TGGATTGATGTCACGATGCTCTGCTTCATCTCTGGAACGCTTGGGAAGACCGAGAA 12410
Qy 1261 acagatgaagtcgctgctgatagggtccgtgatgacctccaccagactcaattttcaacgag 1320
Db 12409 ACAGATGAAGTCGTCGTATAGGTCCTGTATGATCTCCACAGACTCAATTTTCAACGAG 12350
Qy 1321 tctgacgagaatctcaagagtgcctgtctgtaaatccgctgaaatctcctaaacccggtgaa 1380
Db 12349 TCTGACGAGAATCTCAAGAGTGTCTGTCTGAAATCCGCCCTGAATCTCAAAACCCGTTAA 12290
Qy 1381 tcaactccgctcgcgctcatctcaacgaagatacaacgaatcaacacctcgaacgagcgaatg 1440
Db 12289 TCAACTCCGCTCCGATCATCTCCAAACGAAGATCAACAACTCAACCTCGAAGCAGGGGATG 12230
Qy 1441 gtcaacagaacatgctcggcgctaaaccaaactcgcttacttggctttagcgcgagcgcg 1500
Db 12229 GTCAACAGAAATGCTCGCGCGTAACCAAAATTCGCTTACTTGGCTTTAGCCGAGCGC 12170
Qy 1501 tggcttaagttcaggatttcgctaaagtgtgatctcactactgagaaagttaagaacat 1560
Db 12169 TGGCTTAAAGTCTCAGGATTCGCTTAAAGTTGATCTCACTACTCGAGAAATTAAGAAACAT 12110
Qy 1561 ctttacgcgataaccgttcaggagagagcctctgtttctcccgagagaagaggagag 1620
Db 12109 CTTTACGCGGATAACCGTTACGGAGGAGAGCCTCTGTTTCTCCCGGAGAGAGGAGAG 12050
Qy 1621 gaagcagagagatacatcctctgttcttcacgacgagagacatggaatcggagtta 1680
Db 12049 GAAGACGAGGATACATCTCTGTTTCGTTTACGACGAGAGACATGGAATCGGAGTTA 11990
Qy 1681 cagatagttaacgcgcttagcttagaggtttgaagcaacggtttaaacttcctcaggggtt 1740
Db 11989 CAGATAGTTAAACCGCGTTAGCTTAGAGGTTGAAGCAACCGGTTAAACTTCCGTCAGGGGTT 11930
Qy 1741 ccgtacgatttcacggtacattcatcggacgcgagatgattggcgaaacaggtcgtgtga 1800
Db 11929 CCGTACCGGATTTCACGGTACATTTTCATCGGACCGGATGATTGGCGAAGCAGGTCGTGTGA 11870

RESULT 4
LOCUS AY056255
DEFINITION Arabidopsis thaliana putative 9-cis-epoxycarotenoid dioxygenase (MOA2.4/AT3g14440) mRNA, complete cds.
ACCESSION AY056255
VERSION AY056255.1 GI:15810432
KEYWORDS FLI CDNA
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
```

REFERENCE  
AUTHORS

Rosidae; eusoids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 2331)  
Yamada, K., Banh, J., Banno, F., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayaishiaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Kosema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tracy, S.E., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Full length cDNA of gene MOA2.4/AT3g14440 (GI:11994214) Unpublished

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

2 (bases 1 to 2331)  
Yamada, K., Banh, J., Banno, F., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayaishiaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Kosema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tracy, S.E., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Direct Submission  
Submitted (12-SEP-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayaishiaki, Y. and Shinozaki, K.

TITLE  
JOURNAL  
COMMENT

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Banh, J., Banno, F., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Kosema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tracy, S.E., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

FEATURES  
source

1..2331  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="RAFL08-11-H16 (R11094)"  
/notes="This clone is in a modified pBluescript vector (pUC-1) as a BamHI/XhoI insert."  
ecotype: Columbia"

## gene

1..2331  
/gene="MOA2.4/AT3g14440"  
1..122  
/gene="MOA2.4/AT3g14440"  
123..1922  
/gene="MOA2.4/AT3g14440"  
/codon\_start=1  
/evidence=experimental  
/product="putative 9-cis-epoxycarotenoid dioxygenase"  
/protein\_id="AAL07104.1"  
/db\_xref="GI:15810433"

## 5'UTR

## CDS

/translation="MASPTATAVSGRWLGGNHTOPPLASSQSSDLSYSCSLPMASRPV  
TRKLNVSALHTFPALHPFKQSSNSPAIVVKPKAKESNTYQMNLQFRAAAALDAARG  
FLVSHKHLPLPKTADPSVOIAGNFAVNEQPVRRNLPPVVKLPDSISKGVYRNGANP  
LHEPVTGHHFFDGDGMVHAVFPHGSASACRFQTNRFVQERQLSPVFPKPAIGELH  
GHTGIARLMFLYARAAGIVDPAHGTGVANAGLVYFNGLLANSDDLPTQVQITPNG  
DLKTVGRFDPDGQLESTMIAPKVPDPSGELFALSVDVVKYKFRFPDGTKSPD  
VEIQLOQPTMMDHFAITENFVVVDDQVVKLPKPEMILGSGSPVVDKNKVARFGLDKY  
AEDSNIKWIDADPCFCFLHNAWEFETDEVVVGSCMTTPDPSIENESDENLKSVL  
EIKNLKUTESTRRPILISNDEQQVNLNLEAGVNRNMLGRKTFAYLALAEWPVKVSGFA  
KVDLTATGCVKRLHYGDNRGEPFLFPGGEGDEGYILCFVHDEKWTWSELOIVNAV  
SLEAVTVKLPSPVPGFHGTFTIGADDLAKQVV"

```
misc_difference 1234
/ gene="WOA2.4/AT3g14440"
/ note="compared to genomic sequence resulting in an amino
acid sequence difference"
/ replace="g"
1923 - 2331
/ gene="WOA2.4/AT3g14440"
misc_difference 2059
/ gene="WOA2.4/AT3g14440"
/ note="not present in genomic sequence"
misc_difference 2312
/ gene="WOA2.4/AT3g14440"
/ note="not present in genomic sequence"
misc_difference 2315
/ gene="WOA2.4/AT3g14440"
/ note="not present in genomic sequence"
BASE COUNT 615 a 560 c 533 g 623 t
ORIGIN

Query Match 99.9%; Score 1798.4; DB 8; Length 2331;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1799; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 atcggtttttcacggcaacgctcggtttcttgagatggttggtggcaatacatact 60
Db 123 ATGGCTCTCTTTCACGGCAACCGCTCGGTTTCTGGGAGATGGCTTGGTGGCAATCATACT 182

Qy 61 cagcgcattatcgtcttctcaagctccgacttgagttattgtagctcttacctatg 120
Db 183 CAGCCGCCATTATCGTCTCTCAAGCTCCGACTTGAGTTATTGTAGCTCTTACCTATG 242

Qy 121 gccagtctgtcacacgtaagctcaatgtttcatctgcgcttcacactctccagctctt 180
Db 243 GCCAGTCGTGTCACAGTAAGCTCAATGTTTCATCTCGCTTCACACTCTCTCCAGCTCTT 302

Qy 181 catcttcctaagaacatcacaactctccgccattgttgtaagcccaaaagaa 240
Db 303 CATTTCCCTTAAGCAATCATCAAACTCTCCGCCATTGTTGTTAAAGCCCAAGCCAAAGAA 362

Qy 241 tccaactaaacagatgaattgttccagagagggcgggcgagcgttgagcggcg 300
Db 363 TCCAACTAAACAGATGAATTTGTTCCAGAGCGCGCGCGCGGCTGACCGCGG 422

Qy 301 gaggtttctctgtcagccacgagaagctacaccgcttctctaaacggctgactctagt 360
Db 423 GAGGTTTCTTGTGAGCCACAGAGAGCTACACCGCTTCTTAAACGGCTGATCTTAGT 482

Qy 361 gttcagatgcgcggaattttgtccggtgaatgaacagccgctccggcgtaattcttcg 420
Db 483 GTTCAGATCGCCGGAATTTTGTCTCCGGTGAATGAACAGCCGCTCCGGCGTAATCTTCG 542

Qy 421 gtggctcgaataacttccgattccatcaagagtgatgtgcgaacgagactaaccca 480
Db 543 GTGGTCGGAATAACTTCCCGATTTCATCAAGAGAGTGATGTGCGAACGGAGCTAACCCA 602

Qy 481 ettcagagccggtgacaggtcaccactcttcgacgagagcggatggtttcacccgc 540
Db 603 CTTACAGCCCGGTGACAGGTACCACTTCTTCGACGGAGAGCGTATGTTTACCCCGTC 662

Qy 541 aaattcgaacacggttcagctagctacgcttccgggtttactcagactaacgggtttgt 600
Db 663 AAATTCGAACCGGTTACGTAGCTAGCTTACGCTTCCCGTTTACTCAGACTAACCGGTTGTT 722

Qy 601 caggaaactcaattgggtcagccggtttttcccaagccatcggtgagcttcacggccac 660
Db 723 CAGGAACGTCAATTTGGGTGCACCGGTTTTCCTCCAAAGCCATCGGTGAGCTTCACGGCCAC 782

Qy 661 accggtattgcccactcatgtattctacgccagagctgcagccggtatagtcgaccgc 720
Db 783 ACCGTTATTGCCGACTCATGTATTCTACGCCAGAGCTGCAGCCGCTATAGTCACCCG 842

Qy 721 gcacaggaacccggtgtagctaaccggttttggtctatttcaatggcggttattggtct 780
```

RESULT 5

```
Db 843 GCACACGGAACCGGTGTAGCTAACGCCGTTTGGTCTATTTCATTCGCCGTTATTGGCT 902
Qy 781 atgctcggaggatgatttaccttaccagtccagatcactcccaatggagatttaaaacc 840
Db 903 ATGTCGGAGGATGATTTACCTTTACCAAGTTTCAGATCACTCCCAATGGAGATTTAAAC 962
Qy 841 gttgctcgggttcgattttgatggacaattagatccacaatgatggccaccgcgaagtc 900
Db 963 GTTGTGCGTTTCGATTTTGTATGGACAAATTAGAAATCCACAATGATTCGCCACCGAAAGTC 1022
Qy 901 gacccggaatccggtgaactcttcgctttaagctacgacgtcggtttcaagccttaccta 960
Db 1023 GACCCGGAATCCGGTGAATCTCTTCGCTTTAAGCTACGACGTCTTTCAAAGCCTTACCTTA 1082
Qy 961 aaatacttccgattctcaccggacggaactaaatcacggagctcgagattcagcttgat 1020
Db 1083 AAATACTTCCGATTTCTACCGGACGGAATAAATCACCGACGTGAGATTCAGCTTGAT 1142
Qy 1021 cagccaacgatgatgcacgatttcgcgattacagagaacttcgctcgtacctgaccag 1080
Db 1143 CAGCCAACGATGATGCACGATTTCCGATTTACAGAGAACTTCGTCGTCTGCTACCTGACCA 1202
Qy 1081 caagtcgttttcaagctgccggagatgatccggtgggttcctccggttccttcacacaag 1140
Db 1203 CAAAGTCGTTTCAAGCTGCCGAGATGATCTCTCGTGGGTCTCCGGTGGTTTACGACAA 1262
Qy 1141 aacaaggtcgaagattcgggattttagacaataacgccaagatttcacgaacttaag 1200
Db 1263 AACAAAGTTCGCAAGATTCGGGATTTTAGACAATACGCCGAAGATTCATCGAACATTAAG 1322
Qy 1201 tggattgatgtccagattgcttctgcttccatctctggaaacgcttgggaagaccagaa 1260
Db 1323 TGGATTGATGCTCCAGATTGCTTCTGCTTCCATCTCTGGAACGCTTGGGAAGACAGAA 1382
Qy 1261 acagatgaagtcgctgctgatatagggtcctgctgctgctgctgctgctgctgctgctg 1320
Db 1383 ACAGATGAAGTCGTCGTGATAGGGTCTCTGATGATCTCCACAGACTCAATTTTCAACGAG 1442
Qy 1321 tctgcaggaatctcaagagtcctgctgctgctgctgctgctgctgctgctgctgctgctg 1380
Db 1443 TCTGACGAGAATCTCAAGAGTGTCTCTGCTGTAATTCGCCCTGAAATCTCAAAACCCGTA 1502
Qy 1381 tcaactcggctcggatcatctccaacgaagatcaacgaagcaacccctggaagcgggag 1440
Db 1503 TCAACTCGCGCTCCGATCATCTCCAACGAGATCAACAGTCAACCTCGAACAGGGATG 1562
Qy 1441 gtaacagaaacatgctcggcgctaaacacaaatctcgttacttggcttttagccgagccg 1500
Db 1563 GTCACAGAAAACATGCTCGCGCTAAAACCAAAATTCGCTTACTTGGCTTTAGCCGAGCG 1622
Qy 1501 tggctaaagtctcaggattcgctaaagtctgctcactactggagaagtaagaacat 1560
Db 1623 TGGCTAAAGTCTCAGGATTCGTAAGTTCATCTACTTGGAGAAGTTAAGAAAACAT 1682
Qy 1561 ctttacggcgataacccgttacggagagagcctctgcttctcccgagagagagagag 1620
Db 1683 CTTTACGCGGATNACCGTTACGGAGAGAGCTCTGTTTCTCCCGGAGAGAGAGAGAG 1742
Qy 1621 gaagacgaagatacatcctctgtttcttcacgacgagaagacatggaatcggagttta 1680
Db 1743 GAAGACGAAGGATACATCTCTGTTTCGTTCAACGAGAGACATGGAATTCGGAGTTA 1802
Qy 1681 cagatagtaacgcccgttagcttagaggttgaagcaacgggttaaaactccgtcgaagggt 1740
Db 1803 CAGATAGTTAAACGCCGTTAGCTTAGAGTTGAAGCAACCGGTTAAATCTCCGTCACAGGG 1862
Qy 1741 ccgtacggatttcacggtacattcatcgagccggtatgttgccgaagcaggtcgtgga 1800
Db 1863 CCGTACGGATTTACGGTACATTTATTCGAGCCGATGATTTGGCGAAGCAGGTCGTGTGA 1922
```

AC013430	AC013430	95769 bp	DNA	linear	PLN 26-MAY-2000							
LOCUS	Genomic sequence for Arabidopsis thaliana BAC F3F9 from chromosome 1, complete sequence.											
DEFINITION												
ACCESSION	AC013430											
VERSION	AC013430.5	GI:8096768										
KEYWORDS	HTG.											
SOURCE	thale cress.											
ORGANISM	Arabidopsis thaliana											
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.											
	1 (bases 1 to 95769)											
AUTHORS	Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C., Shinn,P., Altafi,H., Bei,Q., Chin,C., Chiou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu.G., Federspiel,N.A., Theologis,A. and Ecker,J.R.											
	Theologis,A. and Ecker,J.R.											
TITLE	Genomic sequence for Arabidopsis thaliana BAC F3F9 from chromosome 1											
JOURNAL	Unpublished											
REFERENCE	2 (bases 1 to 95769)											
AUTHORS	Ecker,J.R.											
TITLE	Direct Submission											
JOURNAL	Submitted (11-NOV-1999) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA											
REFERENCE	3 (bases 1 to 95769)											
AUTHORS	Ecker,J.R.											
TITLE	Direct Submission											
JOURNAL	Submitted (07-FEB-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA											
REFERENCE	4 (bases 1 to 95769)											
AUTHORS	Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bei,B., Chin,C., Chiou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu.G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.											
TITLE	Direct Submission											
JOURNAL	Submitted (24-MAY-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA											
REFERENCE	5 (bases 1 to 95769)											
AUTHORS	Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bei,B., Chin,C., Chiou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu.G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.											
TITLE	Direct Submission											
JOURNAL	Submitted (26-MAY-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA											
COMMENT	On May 26, 2000 this sequence version replaced gi:6921155.											
FEATURES	Location/Qualifiers											
source	1..95769											
	/organism="Arabidopsis thaliana"											
	/db_xref="taxon:3702"											
	/chromosome="1"											
	/clones="F3F9"											
CDS	join(71..402,574..649,733..933)											
	/note="similar to oxidosqualene cyclase dbj BAA33462.1"											
	/codon_start=1											
	/evidence=not_experimental											
	/product="F3F9.1"											
	/protein_id="AAF71792.1"											
	/db_xref="GI:8052528"											
	/translation="MGHVOLLTPCFYSFPFFSFFLLFFLNFNAKNVFLISVDLTVGTV FELEAKVEOVIPIVKVEDGESHTETRRMYRGEFFLSRLCRPMTVGLRSLDLSS FLHWYSSFGHLESIFDAEHREELRIYIYHLNDGGWGLHVEGKSFFCTALNYI CLRILREGDEGREYACKTSKMPWITVLLIYNNECSFFPS1"											
	/note="similar to F-box domain gb AAF14684.1"											
	/codon_start=1											
	/evidence=not_experimental											
	/product="F3F9.2"											
	/protein_id="AAF71813.1"											
	/db_xref="GI:8052549"											
	/translation="MLASWISQTSMKHTKNLILANFNGRFAFTGDNVLKLTETFN HPSRLTSLYOYILODAQAFNCKNLKLDLFOIFAEDVNFNEIVISCSFLEVLILQI IFNPAGHLKIDHKTLRIILSMSCNOJDSIEVRAARDILSIEVIPESDNVILQIPEL QFGRNVWAGRLUPHITSINISCPQKQESGNVMWNTYATSPASISVSLDLKNPREV EVURKILAVWTEKMIIEVTSFKNPNLGEENSDGGAQNKLWEKAEPPFNADPRVDT IMHNFKGSNGKEQFALASRFVMQTKVMKMIKTSFDEKKKKEIEAAVAKLKLPGKN EELSIECF"											
	/join(7138..7400,7491..7887)											
	/note="unknown protein; similar to EST gb AI999747.1"											
	/codon_start=1											
	/evidence=not_experimental											
	/product="F3F9.3"											
	/protein_id="AAF71793.1"											
	/db_xref="GI:8052529"											
	/translation="METGNILKLSLVSVVGYAQAPAPWNPNGPRPGTCDHYE CPTKLVAGYGFIEIRMYDAALMISTPISLMTQATKTGFRRLRYIEGDKSNVK MNMTAPVIAQATPGRSVYTVSYLPPKNNQNPQADDLHVRSTKPTTYAVARVIGGVS NNVAKDEAAALMESLRDSNWILPIEKSGKLPAYFLAVYNPPSHTTARVINEIMVPPN M"											
	/join(8524..8759,8846..9287)											
	/note="hypothetical protein"											
	/codon_start=1											
	/evidence=not_experimental											
	/product="F3F9.4"											
	/protein_id="AAF71794.1"											
	/db_xref="GI:8052530"											
	/translation="MEAAISIFKLSTLLISGSDLIIDSLPNCNRAECPSYEVVHA GNGEYHRNTTWTWISTEPIQDLSNEAGNWNQSLDYNGNNDVHORIEALPIAT QVSNLSITFVFFPKAFQPPPPGNLHVQRWDSRYAVAKQISGVADHKLQKQVA ELKASIQGTWAKAIEKSRGTGGVGSWAYTVAQSFQWQSORVNEINFPFMEDEE TVSIIVQT"											
	/join(13444..13814,13956..14310,14399..14662)											
	/note="similar to dioxxygenase protein"											
	/codon_start=1											
	/evidence=not_experimental											
	/product="F3F9.5"											
	/protein_id="AAF71795.1"											
	/db_xref="GI:8052531"											
	/translation="MAVLSPVAILPKSGFLIPVIDMSPDKHALVKACEDRGGFKV INHGVSAELVSVLEHTVTPFPSPKSEKTQVAGYPGYGNKIGRNGDYGWVEYLLMN ANHDSGGPLFPLSLKSPGTFRNALBEYITSVKMTFDVLEKITDGLIGIKRNTLSKL VQDQNTSILRLNHYPCPLSNKKTNGKNGVIGFGBHTDQIIISVLRNNSTGLQNL NDSQNTSVPPDHTSFPPVNGDSLQVMTNGRFKSVHRVLANKCKRSVSMIYFAGPSLT QRIAPLCLIDNEDELYEEFTWSEYKNSTYNSRLSDNRLOQPERKTIKNLLN"											
	/complement(join(15755..16715,16810..16829))											
	/note="hypothetical protein"											
	/codon_start=1											
	/evidence=not_experimental											
	/product="F3F9.6"											
	/protein_id="AAF71812.1"											
	/db_xref="GI:8052548"											
	/translation="MPKPLRGTSBELPQRSPRLTSLTSSDPHLLSRPITDRSP KGLDRSRSPRGPHDTPUSKGLGRIISGSLQGAQEEELRLKQQLAKAAEAKKR AQEELHKRSKKNTPAPERDDITPGDHQETDVFVLEDEKASEKTKFNDELSAKEDQ INVLKARLYDLKERVLSBENETLDQLKKTDTMSKAKADEKATSKVSOIQEELE ESNETTKLKKLESVEEAKETLEAMKKLVQTEQWKAADAAAALVSGVEMNGRF SEQGSMEXHFAGRFVSGMADDDSDGSGKRSKSGKWFCDLWKKGQK"											
	/complement(join(19162..19689,19767..19833,20044..20164,20234..20378,20482..20623,20693..20738,20826..20877,20965..21156))											
	/note="unknown protein; similar to ESTs gb T41672.1, gb AI992710.1, and gb AA586078.1"											

```
/codon_start=1
/evidence=not_experimental
/product="F3f9.7"
/protein_id="AAF71811.1"
/db_xref="GI:8052547"
/translation="MGKLGKRVVVEERYTKPOGLYVYNKDVVKLKLKLVESKLAP
CYPGDDECHDLEPCIFELYYPSLNRRCCMKSCIECFPLQKMPNSARTQKPLIN
LMGFFPVSLLTFMPCPKCTPNAYEYRGVKSKEEKIEQVBEORVIEAKIRMQKE
MODDERKORLESSESTAMTEMYEGSTGFLVFWFAAISYNSLMDGDEAPSQ
NASVYORSRGRNEDEVDLELMYMEALWLSVQETGTQRNSASGEITSSRYVT
DNHSYVSPRPVPTVEPATPSGGLSCALSALEAOMYGESSHNNHNNVNVSSY
SMLPCNSYVDIEOEVDGIDNHHHRHHYEMGETGSSNSYVSSYMTGEGHNPPPP
PLVIVPSEFEGOMMAVSAEVAHATTTCAPTEVTWQ"
complement(join(2808..23063,23118..23244,23511..23706))
/feature="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F3f9.8"
/protein_id="AAF71810.1"
/db_xref="GI:8052546"
/translation="MSGRGKVKSEPMKVVFINTQYVETDARSFKTVVQELTGKNAIVA
AGPDSPAFDRCYDGGSKIGPITNVVFWKIMSYLSVRPMSNPKRAVYSFHLFK
RTMKPTNYONKCKBCVTKALKATSLRSQASTMGFLWGVPATQVQLGIFTFISFEE
VLYFWFFESCIMELTKLILSLVLKLV"
join(26752..26907,27020..27871,27923..28153)
/feature="similar to expopolygalacturonase precursor
sp|Q00293|PGLX_ASPTU"
/codon_start=1
/evidence=not_experimental
/product="F3f9.9"
/protein_id="AAF71796.1"
/db_xref="GI:8052532"
/translation="MVSTIGSFLIFIAVASSISAAPSAAALVGRKVFDRSYGARG
DKTNDTMAFTKAMQKQWGLPRVYIPFGTFYLGAVAFTPCKSRISFTIKGTLIA
PKDPAIKQDSMIIPRYVDYLVSGGILDGGSYSWPLNNCROTHNCRLPLMNGFO
FVRSRLRIKISINSMGHLAFPSVQHEFDITRVNIKAPGDSPTDGIKIGSNHMKIF
HVDIATDDCIALLSGTNLDINKVNCFGHGHSVGLKFKGKSVGLLVIRNSIFN
GTSNGVRIKTVKPSPEPLNVNLFPLQIMQIDQYVFNIDQYCPNPPCSFOVTSLTR
NKFSKIQIRVFKFQNIWGTSTAKEAVLQCSKNVPCKNVQLFNINIVHRGRDGPATS
VCENVGWGGIKSPSPSCIR"
33566..35539
/feature="similar to 9-cis-epoxycarotenoid dioxygenase
gp|AAF26356.1"
/codon_start=1
/evidence=not_experimental
/product="F3f9.10"
/protein_id="AAF71797.1"
/db_xref="GI:8052533"
/translation="MTLITISGMVYISLLSDAHHSQYQNTNLVLKPKIPQTAIA
FNQESTMASTLLPSTQFLDRFTSTSSSRPKLQSLFSSTLRNKKLVVPCYVSS
SVNKKSSVSSLSQSPFTFKPSPKLCNDVNLNPKTNNQKLPVORVTAAMVLDVAE
NMLSHERRRHPKPTADPAVQIAGNFPVPEKPVHNLVPTGTVPICIQGVYVRNGA
NPLHKPVSGHILFDGDMVHVRDNGSVACRFTENRLVQRECEGRPVFPAIGE
LHGLGIKLMPLNTRGLVGLVPTGGLGVANAGLVYFNGHLLAMSEDDLPYHVKTQ
TGDLETSRYDFDGLQKSLMIAHPKIDPETRELFALSYDVVSKPYLYFRFTSDGKS"
Query Match 47.9%; Score 862.8; DB 8; Length 95769;
Best Local Similarity 73.1%; Pred. No. 2.1e-245;
Matches 1139; Conservative 0; Mismatches 407; Indels 12; Gaps 2;
QY 245 acactaaacagatgaattgttcagagagcgcgcgagcgatgttgagcgcgaggg 304
Db 33987 AGAATCCAAAATTAACCCGGTGCAGAGAACCCGGCTATGCTTTTAGACCGCGGTGAAA 34046
QY 305 gtttcctgtcagcacag---aagctacacccgtctcctaaacgctgactagtq 361
Db 34047 AGCCATGATCTCACACGAGCGCGCGTTCATCCCCCATCTTAAACCGCGGATCTCGCG 34106
QY 362 ttcagatcccggaatttgcctcgggtgaatgaacagccgctccggcgtaattcttcgg 421
Db 34107 TCAATAATAGCCGGGAATCTTCTCCGGTACCGGAGAAACCTGTCTGATACCTTCGG 34166
QY 422 tggctcgaaacttccccgattccatcaaaaggagtgatgtgcgaacggagctaacccac 481
```

```
Db 34167 TGACTGGAAACAGTACCAGAAATGCATTCAAGGAGTTTTCAGTACAGAACGGAGCAATCCAC 34226
QY 482 ttcacgagccgggtgacaggtccaccactcttcgacgggagacgggtatgggtcacgcgctca 541
Db 34227 TTCACAAACACAGTCTCCGGCCACCAATTTGTTGCGAGGAGACGGTATGGTACAGCTGTCC 34286
QY 542 aattcgaacacgggttcagctagctacgcttgcggtttactcagactaacacgggtttgttc 601
Db 34287 GGTTCGATAAACGGCTCGGTAGCTACGCTTGTGCGTTTACCGGAAACAAACCGGTGGTTC 34346
QY 602 aggaacgctcaattgggtcgacccgggtttccccaaagccatcggtgagcttcacggccaca 661
Db 34347 AAGAAACGTGAATGTGTGCTCGTCCGGTTTCCCAAGAGCAATTCGAGAGCTTCACGGACAT 34406
QY 662 ccggtattgcgcgactcatgctattctacgcccagagctgcagccggtatagtcgaccccg 721
Db 34407 TAGGAATGCCCAAGCTTATGCTTCTCAATACCCGGGGCTATTTGGTTAGTCGACCCGA 34466
QY 722 cacacggaacgggtgtagctaacgcgggtttgggtctatttcaatggcgggttattggcta 781
Db 34467 CCGGAGGACTCGGTGCTGCTAACCGCGGTTTAGTTTATTTCAATGGTCATCTCTTAGCCA 34526
QY 782 tctcggagagatgatttaccttaccaggttcagatcactcccaatggagattttaaaaccg 841
Db 34527 TGTCCGAAAGACGATTTACCGTACCATGTCAAAGTTTACTCAAACCGGAGATTTAGAAACT 34586
QY 842 ttggttcgggttcgattttgatggacaatttagaattcaccaatgattccccaccggaagtcg 901
Db 34587 CGGTGCGGTACGATTTGACGGTCAGTTAAATCAACAAATGATAGCCCAACCCGAAATCG 34646
QY 902 acccggaatccgggtgaactcttcgctttaagctacgagctgctttcaagccttaccctaa 961
Db 34647 ATCCGGAACCCGAGAACTTTCGCTTTTAAGCTACGACGTCGCTTTCAAAGCCTTACTTAA 34706
QY 962 ataacttcgattctcacggagcgaactaataaccggagcgtcagatcagcttgatc 1021
Db 34707 AATACTTCAGATTACATCCGACGGTGAGAAATCAACAGACGTCGAGATTCGCTTGATC 34766
QY 1022 agccaacgagatgacgatttcgcgattacagaaacttcgctcgtacctgaccgaccg 1081
Db 34767 AACCAACGATGATCCACGATTTCCGATACCTGAGAAATTTCTAGTAGTTCAGACCAAC 34826
QY 1082 aagtcggtttcaagctgcggagatgacgagtcggttgggtctcgggtggttttacgacaaag 1141
Db 34827 AAGTGGTATTCGATTGCGGAGATGATAAGAGGTGGTTCCTCCGCTGGTTTACGACGAGA 34886
QY 1142 acaaggtcgcaagattcgggatttttagacaataacgcgcaagatttcgaacattaagt 1201
Db 34887 AGAAGAAATCAAGATTCCGGGATTTTGAATAAAACCGGAAAGATGCTTCGTCGATTCAAT 34946
QY 1202 ggattgactccagattgcttcttcctcctcctcctcctcctcctcctcctcctcctcctcct 1261
Db 34947 GGATCGAAGTACCAGATTGTTTTTTTTCATCTATGAAACTCTTGGGAAAGAACAGAGAA 35006
QY 1262 cagatgaagtcgctgctgatagggtcctgatgactccaccagactcaatttttcaacgagt 1321
Db 35007 CAGACGAGTTTGTGCTGATGGATCATGATCAGCCACCTGATTCAATTTTCAACGAAAC 35066
QY 1322 ctgacgagaatctcaagagtgctcctgtctgaataccgctcctgaatctcaaacctgagat 1381
Db 35067 ACAGCAAAACACTTTCAGAGTGTGTTTGTGCGAGATAAGAGCTAAACCTTAAACACAGGG 35126
QY 1382 caactcgcgcgctcagatcatctcccaacgaagatcaacagctcaactcgaacgagatgag 1441
Db 35127 CAACACGTAGACCGGTTATCTC-----GGAAACAGTTAAATCTCGAAGCTGTGATGG 35177
QY 1442 tcaacagaaacatgctcgccgctgtaaaacaaactcgttacttggctttagccggagcgt 1501
Db 35178 TCAACCGGAAATTTATTAGTAGAAAAACCCGGTATGCTTATCTGCTTTTAAACCGAACCGT 35237
QY 1502 ggcctaaagcttcaggattcgtctaaagttgattcactactcggagaggttaagaacatc 1561
```

Db 35238 GGCCTAAAGTGTCCGGTTTCGCTAAAGTGGACTTATCGACCGGAGAGAGATTCCGAAGTATA 35297  
Qy 1562 ttacggcgataaacggttacggagagagagcctctgttcttctcccccgaagaagagagg 1621  
Db 35298 TTTACGGAGAGGGGAAATACGGAGGAGAGCCTCTGTTTCTACCTTCGCGTACGGAGAAG 35357  
Qy 1622 aagcgaaggaatacatcctctgttccgttcacgacgagaagaacatggaaatcggagttac 1681  
Db 35358 AAGACGGAGGTTACATAAATGGTGTTCGTTTCACGACGAGAGAGAGTGAATTCGGAACCTC 35417  
Qy 1682 agatagtttaacgcccgttagcttagaggttgaagcaacggttaaaccttccgtcaagggttc 1741  
Db 35418 AACTATAACGCCGTTTAATATGAAGCTTGAAGCTACCGTAAACGCTTCGTCGAGAGTCC 35477  
Qy 1742 cgtacggatttcaacgggtacatcatcggagccgatgatttggcgaagcaggtcgtgtg 1799  
Db 35478 CTTACCGGTTTCACGGAACGTTTCATCAGTAAGGAAGATTTCATCGAAGCAAGCTTTGTG 35535

RESULT 6  
AX148316 AX148316 1818 bp DNA linear PAT 08-JUN-2001  
LOCUS  
DEFINITION Sequence 15 from Patent EP1116794.  
ACCESSION AX148316  
VERSION AX148316.1 GI:14347203  
KEYWORDS  
SOURCE tomato.  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.  
REFERENCE 1 (bases 1 to 1818)  
AUTHORS Iuchi,S., Kobayashi,M. and Shinozaki,K.  
TITLE Transgenic plants carrying neoxanthin cleavage enzyme gene  
JOURNAL Patent: EP 1116794-A 15 18-JUL-2001;  
Riken (JP)  
FEATURES  
source Location/Qualifiers  
1..1818  
/organism="Lycopersicon esculentum"  
/db\_xref="taxon:4081"  
1..1818  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC41199.1"  
/db\_xref="GI:14347204"  
/translation="MAITTSHTATNTWIKLSMPSEKFGFASNSISLLKNQHNRSQSL  
NINGSLOAPPLIHPKOSNSYOTPKNNTISHPKOENNNSSSSSTSKWNLVQKAAMAL  
DAVSALTKEHLEHLPKTDPRVQISGNFAPVPENVCOSLPVTGKI PKCVQGVYVR  
NGANPLEPTAGHHFFDGDGMVHAVQPKNGSASACRPTERLVQEKALGRPVFPKA  
IGELHSGIARLMFLFARGLFGVLDHSGTGVANAGLVVFNNRLLAWSBDDLLPYHK  
VTPTGDLKTEGRFPDQOLKSTMTIAHPKLDVPSGELFALSDYVIQKPYLYKFFRSKNG  
EKSNDVEIPVBDPTMHDFAITENFVVI PQQVVFKNSEMI RGSPVYVDSIKNKYSRFG  
ILDVAKGDSGLKLVWPDPCFPLMNAWEAEDEIWIWISGMPDPSI FNECDEGL  
KSVLEIRLNLTKCKSTRKSIENPDQVNLNLEAGVNRNKRKTEYAVILAIAPWPVK  
VSFAKYNLFGEVEKFIYCDNKYGGEPFLPRDPNSKEEDDGYLLAFVHDEKENKSE  
LQIVNMSLKLEATVLPFSRVPYFGHGFIFINANDLANQA"  
BASE COUNT 569 a 383 c 369 g 497 t  
ORIGIN

Query Match 45.8%; Score 824.6; DB 6; Length 1818;  
Best Local Similarity 69.6%; Pred. No. 2.5e-284;  
Matches 1134; Conservative 0; Mismatches 489; Indels 6; Gaps 1;  
Qy 170 ctcagctcttcttcccttaagcaatcatcaactctcccgccattgtgttaagccca 229  
Db 185 CTTCAAAATTATCAACACCAAGAATAATACAAATTCACCCCAACCAAGAAACACA 244  
Qy 230 aagccaaagaatcccaacataacagatgaattttgttccagagcggcggcagcgt 289  
Db 245 ACTCTCTCTCTTCAACTTCCAAGTGGAAATTTAGTGCAGAAAGCAGCAGCAATGGCTT 304

Qy 290 tggcgcggcgaggggttccttgtcagccacgagaagctacacccgcttctctaaacgg 349  
Db 305 TAGATGCTGTAGAAAGTGTCTTTAACTAAACATGAACCTTTGCGGAAACAG 364  
Qy 350 ctgacccatagtttcagatcgccggaattttgctccgggtgaatgaacacgcccgtccggc 409  
Db 365 CCGACCCACGAGTCCAGATTTCCTGGGAATTTTGTCTCCGGTACCGGAAAAATCCAGTCTGTC 424  
Qy 410 gtaatcttcgggtgggtcggaatacttcccgattccatcaaaaggagtgatgtgcgcaacg 469  
Db 425 AATCTCTTCGGGTACCGGAAAAATACCCAAATGTGTTCAGAGCGCTTACGTTCGAAACG 484  
Qy 470 gagctaaccacacttcacgacgggtgacaggtcacacttcttcgacggagacggtatgg 529  
Db 485 GAGCTAACCTCTCTTTTGAACCAACCGCGACACCAATTTCTTCGACGGGACCGGTATGG 544  
Qy 530 ttcacgcgctgaattcgaacacgggttcagctagctacgcttgcggttttactcagacta 589  
Db 545 TTCACGCGCTTCAATTCAAAAAATGGGTGCGCTAGTTACGCTTGCCTTTCACTGAACAG 604  
Qy 590 accggtttgttcaggaacgtcaattgggtcgacgggttttccccaagccatcggtgagc 649  
Db 605 AGAGGCTTGTTCAGAAAAAGGCTTTGGGTGCGCCCTGTTTTCCCTAAAGCCATTGGTGAAT 664  
Qy 650 ttcacgcccacacgggtatgcccgaactcatgtattctacgcgcagagctgcagccggt 709  
Db 665 TACATGCTACTCTGGAATTCGAAGGCTTATGCTGTTTTACGCTCGTGGGCTCTTCGGAC 724  
Qy 710 tagtcgaccggcgacacgggaacgggtgtagtaaacgcgggtttgggtctatttcaatggcc 769  
Db 725 TTGTTGATCACAGTAAAGGAACCTGTTGTCAAACGCGGGTTTAGTCTATTTCAATAACC 784  
Qy 770 ggttattggctatgctcgagagatgattacccttaccgaagttcagactcactcccaatggag 829  
Db 785 GATTACTTGTCTATGCTGAAGATGATTGGCTTACCATGTAAAGGTAAACACCCACCGCG 844  
Qy 830 atttaaaacggttgggtcggttcgattttgatggacaattagaatccacaatgattcccc 889  
Db 845 ATCTTAAACACAGGGTTCGATTTCGACGGCCAGCTTAAATTCACCATGATAGTCTC 904  
Qy 890 accggaagtcgacccggaaatccgggtgaactcttcgctttaaagtaacgaagtcgtttcaa 949  
Db 905 ACCCAAGCTCGACCCAGTTTCGCGTGAGCTATTTCCTTTAGTCTACGATGTGATTAGA 964  
Qy 950 agccttacctaaataacttccgattctcacggcggaacataatcacggagacgtcgaga 1009  
Db 965 AGCCATACCTCAGTACTTTCAGATTTTCAAAAAATGGGAAAAATCAAAATGATGTTGAAA 1024  
Qy 1010 ttcagcttgatcaccccaacgatgatgcacgatttcgagattcacagaaacttcgtcgtcg 1069  
Db 1025 TTCAGTTGAAGACCCCAACAAATGATGATGTTTCGCAATTTACTTGAGAACTTCGTCGTCA 1084  
Qy 1070 tacctgaccagcaagtcggttttcaagctgcgggagatgatccgcggtggggtctccggtgg 1129  
Db 1085 TTCCTGATCAACAAGTCGTTTTTCAAGATGTCGAAATGATCCGTCGGAGTTTCAACCGGTG 1144  
Qy 1130 tttcacgaagaacaaggctcgagaagattcgggatttttagacaaataacgcgaagattcat 1189  
Db 1145 TTTACGACAGAAACAAAGTTTCCCGGATTTGGTATTTCTGGATAACGCGAAGATGGGT 1204  
Qy 1190 cgaacattaaagtgattgatgctccagatcgcttctctctctctctctgaacgcttggg 1249  
Db 1205 CTGATTTGAATGGTTGAAGTACTGATGTTTCTGTTTCCACTCTGGAAATCGTTGGG 1264  
Qy 1250 aagggcagaaacagatgaagtcgctgatagggtcgtgatgactccaccagactcaa 1309  
Db 1265 AAGAAGCAGAAACAGATGAAATCGTTGTAATTTGTTTCATGTATGACACCCACGACTCCA 1324  
Qy 1310 ttttcaacaggtctgacgagaatctcaagagtgctcgtgctgaaatccgctcgtgaattcca 1369  
Db 1325 TTTTCAATGAATGTGATCAAGGGCTAAAGAGTGTGTTTTATCCGAAATCCGCTCAATTTGA 1384  
Qy 1370 aaacgggtgaatcaactcgcgcgtccgatcatcttcaacgaagatcaaacgctcgaactcg 1429

```

Db 1385 AAACAGGGAATCAACAGAAATCATATATCGAAACCCGGATGAACAGTGAATTTAG 1444
Qy 1430 aagcaggatggtcaacagaaacatctcgccgctgaaacacaaattcgcttacttgctt 1489
Db 1445 AAGCTGGAATGGTGAACCGAAACAAACTCGGAAGGAAACACAGATATGCTTATTTGGCTA 1504
Qy 1490 tagccagcgtggtcctaaagctcaggtatcgctaaagttagctcaactcactgagagag 1549
Db 1505 TCGCTGAACCATGGCCCAAAAGTTCTGGTGTTCGAAAGTAAACCTGTTTCAACGGTGAAG 1564
Qy 1550 taaagaaacatcttacgcgcataaccgttcagggagagcctctgtttcccccgaag 1609
Db 1565 TTGAGAAATTCATTTATGTGTGACAAATAATGTCGGGAACCTCTTTTACCAGAG 1624
Qy 1610 -----aaggaggaggaagcgaagatacatctctgtttcgttcacgcagagaaga 1663
Db 1625 ACCCAACAGCAAGGAAGAGCATGGTGTATATTTTAGCTTTTCGTTACGATGAGAAAG 1684
Qy 1664 catggaatcaggagttacagatagtttaacgcgcttagcttagagttgaagcaacggtta 1723
Db 1685 AATGGAATCAGAACTGCAAAATGTTTAAACGCAATGAGTTTGAAGTTGGAGGCAACTGTGA 1744
Qy 1724 aacttcgcgaaggttcggtacgattccaggtacattcatcggagcggatgattgg 1783
Db 1745 AGCTTCATCAAGAGTTCCTTATGGAATTCATGGAACATTCATAACGCCAATGATTGG 1804
Qy 1784 cgaagcagg 1792
Db 1805 CAATCAGG 1813

RESULT 7
LENEOXANT
LOCUS LENEOXANT 2171 bp mRNA linear PLN 22-MAY-1998
DEFINITION Lycopersicon esculentum mRNA for nine-cis-epoxycarotenoid
dioxigenase.
ACCESSION 297215
VERSION 297215.1 GI:2769641
KEYWORDS nine-cis-epoxycarotenoid dioxigenase.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 2171)
AUTHORS Burbridge,A., Grieve,T.M., Jackson,A., Thompson,A. and Taylor,I.B.
TITLE Structure and expression of a cDNA encoding a putative neoxanthin
cleavage enzyme (NCE) isolated from a wilt-related tomato
(Lycopersicon esculentum Mill.) library
J. Exp. Bot. 47, 2111-2112 (1997)
REFERENCE 2 (bases 1 to 2171)
AUTHORS Burbridge,A.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-1997) Burbridge A., The University of Nottingham,
Physiology and Environmental Science, Sutton Bonington Campus,
Loughborough, Leicestershire, LE12 5RD, UK
REMARK Revised by [3]
REFERENCE 3 (bases 1 to 2171)
AUTHORS Burbridge,A.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-1998) Burbridge A., The University of Nottingham,
Physiology and Environmental Science, Sutton Bonington Campus,
Loughborough, Leicestershire, LE12 5RD, UK
COMMENT On Jan 13, 1998 this sequence version replaced gi:2243153.
FEATURES
Location/Qualifiers
1..2171
/organism="Lycopersicon esculentum"
/db_xref="taxon:4081"
79..1896
/codon_start=1
/product="nine-cis-epoxycarotenoid dioxigenase"
CDS

```

```

/protein_id="CAB10168.1"
/db_xref="GI:2769641"
/translat="MATTSHATNTWTKLSMPSSKEFGFASNSISLLKNOHNRQSL
NINSSLOAPPTLHPKOSNNVTPKNNITSHPKOENNNSSSSSTSKNVLVOKAAMAL
DAVESALTKEHEHPLPKTADPRVQISGNFAPVPENPVCSLPVTGKTPKCVQGVYVR
NGANPLPEPTAGHHFFDQGMVHAFQKNGSAYACRPTETTERLUVQEKALGRVPFPA
IGELHSGIARLMFLFYARGLFGVDSKGTGVANAGLVYFNNLLAMSEDDLPYHKV
VPTGDLKTEGRFDQGLKSTIAHPKLPDVSGLFALSVDVQKPYLTKFRFSKNG
EKSNDVEIPVEDPTMHDFAITENFVVPDOVVFKMSEMRGGSVVYDKNKSVPFG
ILDKVAKGSDLKWVEYDCPCFHLNNAWEAEDELVIIGSCMTPPDSI FNEDEGL
KSLVSELRLNLTKSTKRSIIENPDQVNLNAGVNRNKLGRTEREYALAIAPWPX
VSGPAKVNLTFTGEVKFYDNKYGGSEFLFPLRDPNSKEEDDGYILAFVHDEKEWKSE
LQIVNAMSLEATVPLSPRPYFGHGFNFNANDLANQA"
BASE COUNT 688 a 435 c 431 g 617 t
ORIGIN

Query Match 45.8%; Score 824.6; DB 8; Length 2171;
Best Local Similarity 69.6%; Pred. No. 2.6e-234;
Matches 1134; Conservative 0; Mismatches 489; Indels 6; Gaps 1;

Qy 170 ctccagctcttcatttccttaagcaatcatcaaacctcccgccattgtgttaagccca 229
Db 283 CTTCAAAATTAACAACACCAAGAAATAATACAAATTCACACCCAAAACAAGAAAAACA 322
Qy 230 aagccaaagaatcccaactaaacagatgaattgttccagagagcggcggcagcgt 289
Db 323 ACTCCTCTCTTCTCAACTTCCAAGTGAATTTAGTCGAGAAAGCAGCAATGGCTT 382
Qy 290 tggacgcggcggagggtttcctgtcagccacgagaagctacacccgcttctcctaaacgg 349
Db 383 TAGATGCTGTAGAAAGTGTCTTAACATAACATGAATTCGAACACCTTTTGCAGAAAA 442
Qy 350 ctgacctagtggtcagatcccggaattttgttccggtgaatgaacagccgctccggc 409
Db 443 CGACCCACGAGTCAGATTTCTGGGAATTTTGTCCCGGTACCGGAAATCCAGTCTGTG 502
Qy 410 gtaattcttcgggtggtcggaataactcccgatccattccatcaaaaggagtgatgtgcgcaacg 469
Db 503 AATCTCTTCGGGTACCGGAAAAATACCAATGTGTTCAAGGCGTTTACGTTCCAAAGC 562
Qy 470 gagctaaccttctacagcgggtgacaggtcaccaacttcttcgacggagagcgatgg 529
Db 563 GAGCTAAACCTCTTTTGAACCAACCGCGGACACCAATTTCTTCACGCGGACGCGATGG 622
Qy 530 ttcaacgcgtcaaatcgaacacggttcagctagctagctgacggttcgaggttactcagacta 589
Db 623 TTCACGCGGTTCAATTTCAAAAATGCGGTGCGGTAGTTACGCTTGCCTTTTCACTGAAACAG 682
Qy 590 accggtttggtcaggaacggtcaattgggtcgacgggttttcccaaaagccatcggtagc 649
Db 683 AGAGCTTGTTCAGAAAAAAGCTTGGGTGCGCCCTGTTTTCCCTTAAGCCATTTGTTGAAT 742
Qy 650 ttacggccacacgggtattgcccgaactcatgtattctacgcccagagctgagccggtta 709
Db 743 TACATGGTCACTCTGGAATTCGAAGGCTTATGCTGTTTTTACGCTCGTGGGCTCTTTCGGAC 802
Qy 710 tagtcgacccggcacacggaacccggtgtagtaacgcgggttttggtctatttcaatggcc 769
Db 803 TTGTTGATCACAGTAAGAGAACTGTGTGTGCAAAACGCGCGTTTAGTCTTATTTCAATAACC 862
Qy 770 ggttattggtctatcgaggatgatttaccttaccaggttcagatcactcccaatgag 829
Db 863 GATTACTTGTCTATGCTGGAAGATGATTTGCCCTTACCATGTAAAGGTAAACCCACCGCGG 922
Qy 830 atttaaaacccgttggtcgggttcgattttgatggacaattagaaatcccaatgattgcc 889
Db 923 ATCTTTAAACAGAGGTCGATTTCGATTTTCGACGGCCAGCTAAATCCACCATGATAGCTC 982
Qy 890 accgaaagtcgacccggaatcccggtgaactcttcgcttttaagctacgacgtcggttcaa 949
Db 983 ACCCAAGCTCGACCCAGTTTCCGGTGAGCTATTGTGCTCTTTAGCTACGATGTGATTTCAGA 1042

```



Qy	950	agccttacctaaataacttcgattctccagcgacggaactaaatacccgacgtcaga	1009
Db	1043	AGCCATACCTCAAGTACTTCAGATTTTCAAAAATGGGAAAAATCAAAATGATGTTGAAA	1102
Qy	1010	ttcagcttgatcagccaacgatgatcagagatttcgsgattacagagaacttcgctgcg	1069
Db	1103	TTCCAGTTGAAGACCCAACAAATGATGATGATTTCCGAATTACTGAGAACTTCGTCGTCA	1162
Qy	1070	tacctgaccagcaagtcggttttcaagctgcggagatgatccgcggtgggtctccggtgg	1129
Db	1163	TTCTCTGATCAACAAGTCGTTTCAAGATGCTCAAAATGATCCGTGAGGTTTACCGGTGG	1222
Qy	1130	tttacgacaagaacaggttcgcaagatttcgggatttttagacaataaccgcaagattcat	1189
Db	1223	TTTACGACAAGAACAAAGTTTCCCGATTTCGGATTTCGGATAAGTACCGGAAGATGGGT	1282
Qy	1190	cgaacattaaagtggattgatgtccagattgctctcgtctccatctctggaacgcttggg	1249
Db	1283	CTGATTTGAATGGTTGAAGTACCTGATTTGTTTCTGTTTCCACCTCTCGAATGCTTGGG	1342
Qy	1250	aagagccagaaacagatgaagtcgctcgtgatagggctcctgtatgactccaccagactcaa	1309
Db	1343	AAGAAGCAGAAACAGATGAAATCGTTGTAATTGGTTTCATGTATGACACCAACAGACTCCA	1402
Qy	1310	ttttcaacgagcttcgacgagaactcaagagtgctcctgtctgaaatccgctgaaatctca	1369
Db	1403	TTTTCAATGAATGTGATGAAGGGCTAAAGAGTGTTTTATCCGAAATCCGCTCTCAATTTGA	1462
Qy	1370	aaaccggtgaatcaactcgcgcgcgatcatctccaaagagatcaacaagtcaacctcg	1429
Db	1463	AAACAGGGAATCAACAGAAAAATCCATAATCGAAAAACCCGGATGAAACAAGTGAATTTAG	1522
Qy	1430	aagcaggatggtcaacagaaaacatgctcggcgctgtaaaccaaatctcgcttacttggctt	1489
Db	1523	AAGCTGGAATGGTGAACCGGAAACAAACTCGGAAGGAAAAACAGAGTATGCTTATTTGGCTA	1582
Qy	1490	tagccgagcgcgtggccttaagcttcaggatttcgctaaagttagctcaactactcagagaag	1549
Db	1583	TCGCTGAACCATGCGCCAAAAGTTTCTGGTTTTCGAAAAAGTAAACCTGTTCACCGGTGAAG	1642
Qy	1550	ttaagaaacatctttacggcgataaacggttacggagagagcctctgtttctcccgcgag	1609
Db	1643	TTGGAATTCATTTATGGTGACAAACAATATGGTGGGGAACCTCTTTTTTACCAAGAG	1702
Qy	1610	-----aaggagagagaagacgaaggatacatcctctgtttcgttcacgacgagaaga	1663
Db	1703	ACCCCAACAGCAAGGAAGAAGACGATGGTTATATTTTAGCTTTTCAGTATGAGAAAG	1762
Qy	1664	catggaatcggaggttacaagatagttaagcgcgttagcttagaggttgaagcaacgggta	1723
Db	1763	AATGGAATCAGAACTGCAAAATTTGTTAAACCAATGATGTTGAAGTTGGAGGCAACTGTGA	1822
Qy	1724	aacttcgctcaagggttcgtagcgatttcacggttacattcatcggagcgcgatgttgg	1783
Db	1823	AGCTTCCATCAAGAGTTCTTATGGATTTCATGSAACATTCATAACGCCAATGATTTGG	1882
Qy	1784	cgaagcagg 1792	
Db	1883	CAAAATCAGG 1891	
RESULT 8			
STU276244			
LOCUS			
DEFINITION			
Solanum tuberosum mRNA for putative 9-cis-epoxycarotenoid			
dioxygenase (nced1 gene).			
ACCESSION			
AJ276244			
VERSION			
AJ276244.1 GI:7209268			
KEYWORDS			
9-cis-epoxycarotenoid dioxygenase; nced1 gene.			
SOURCE			
potato.			
ORGANISM			
Solanum tuberosum			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			



```
Db 758 GGGTTATGCTGTTTACGCTGCGGCTCTTCGGAATTAATGATCACAGTAGAGGAACGTG 817
Qy 734 gtgtagtaaacgcgggtttgggtctatttcaatggccgggtatttgggtctatgtcggagatg 793
Db 818 GAGTTGCAAAACGCGGTTTAGTCTATTCAATAACCGATTACTTGTCTGTGAAGATG 877
Qy 794 atttaacttaccgaagtcagatacactcccaatggagatttaaaacccgttggcgttcg 853
Db 878 ATTTCCTTACCATGTAAGGTAAACACCCACCGCGCATCTTAAACAGAGGGTCGATTGC 937
Qy 854 attttgatggacaattagaatcacaaatattgcccaccgaaagtcgaccgggaatccg 913
Db 938 ATTTGACGCGCAGCTAAATCCACCATAGTAGCTCACCAGCTCGACAGCTTTCG 997
Qy 914 gtgaactcttcgctttaagctacgagcgtcggttcaaagccttaccctaaatacttcgat 973
Db 998 GTGAGCTATTTCCTTAGCTACGATGTGATTCAGAAGCCATACCTCAAGTATTCAGAT 1057
Qy 974 tctcaccggacggaactaaatacacggacgtcgagattcagcttggatcagccaagatga 1033
Db 1058 TTTCAAAAAATGGGAAAAATCAAAATGATGTTGAAATCCAGTTGAAGACCCCAACAATGA 1117
Qy 1034 tcacgatttcgagattacagaaacttcgctcgtcgtacactgaccgaaagtcgtttca 1093
Db 1118 TGCATGATTTCGCAATTACTGAGAAATTCGTTCATCAATTCCTGATCAACAAAGTCGTTTCA 1177
Qy 1094 agctgcggagatgatccggtgggtctccggtggtttacacaaagaaagtcgcaa 1153
Db 1178 AGATGCTCGAATGATCCGTGGAGGTTACCCGGTGGTTTACGACAGAACAAAGTGTCCC 1237
Qy 1154 gattcgggattttagacaaataacgcgaagattcatcgaacattaaagtgattgatctc 1213
Db 1238 GATTTGGGATTTCTGATAAGTAGTACCGGAAAGATGGGTCTGATTTGAAATGGGTGAAGTTC 1297
Qy 1214 cgaattgcttcgcttcacatctggaacgcttgggaagcgttgggaagcagaaacagatgagtcg 1273
Db 1298 CTGATTGTTTCTGTTTCCATCTCTGAAATGCTTGGGAAGAACCGAAGAACAGATGAATCG 1357
Qy 1274 tgcgtgataggctcgtatgactccaccagactcaatttcaacgagctctgacgagaatc 1333
Db 1358 TTGTAATTCGTTTCATGATGACACACACAGACTCCATTTTCAATGAATGTGATGAAGGAC 1417
Qy 1334 tcaagagtcctgtctgaaatccgctgaatcctcaaaaccggtgaatcaactcgcgtc 1393
Db 1418 TAAAAAGCGTTTATCCGAAATTCGTCTCAATTTGAAACAGGGAATCAACAAGAAAG 1477
Qy 1394 cgaatcatctccaaagaaatcaacaagtcaactcgaagcaggatggtcaacagaaca 1453
Db 1478 CCATAATCGAAACCCGGATGAACAAAGTGAATTTAGAGCTGGAAATGGTGAACAGAAACA 1537
Qy 1454 tgcctcgccgtaaaaacaaattcgtctacttggctttagccgagccgtggcctaagctc 1513
Db 1538 AACTTGGAGGAAACACAGTAGTGTATTATTTGGCTATTGCTGAACATGGCCAAAGATT 1597
Qy 1514 caggattcgtctaaagtgtactcactcgtggaagagtttaagaacaatcttttcgcgcata 1573
Db 1598 CTGGTTTTCGAAAAGTAGACCTGTTTACAGGTGAAGTTGAGAAAATTCATTTATGGTGACA 1657
Qy 1574 accgttacggagagagcctctgttctctcccgag-----aaggagagagggaagacg 1627
Db 1658 ACAAATATCGTGGGAGCCTCTTTTTTATACCAAGAGACCCAAACAGCAGGAGGAGAGCG 1717
Qy 1628 aaggatacatcctcgttctcgttcacgacgagaagacatcggaatcggaattacagatag 1687
Db 1718 ATGGTTATATTATTAGCTTTCGTTTCATGATGAGAAAGATGGACATCAGAGCTTCAATTTG 1777
Qy 1688 ttaacgcgcttagcttagaggttgaagcaacggtttaaacttcctcgtcaagggttccgtacg 1747
Db 1778 TTAATGCAATGACTTTAAAGTTAGAGGCAACAGTGAACCTTCCATCAAGAGTTCCTTATG 1837
Qy 1748 gatttcaecggtacattcgcgagccgatgattttggcgaagcagg 1792
```

```
Db 1838 GTTTTCATGGAACTTTTCATAATGCAAAATGATTTGGCAAAATCAGG 1882

RESULT 9
LOCUS AX148312 1839 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 11 from Patent EP1116794.
ACCESSION AX148312
VERSION AX148312.1 GI:14347199
KEYWORDS cowpea.
SOURCE Vigna unguiculata
ORGANISM Vigna unguiculata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Vigna.
REFERENCE 1 (bases 1 to 1839)
AUTHORS Iuchi,S., Kobayashi,M. and Shinozaki,K.
TITLE Transgenic plants carrying neoxanthin cleavage enzyme gene
JOURNAL Patent: EP 1116794-A 11 18-JUL-2001;
Riken (JP)
FEATURES
source 1. 1839
/organism="Vigna unguiculata"
/db_xref="taxon:3917"
CDS 1. 1839
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC41197.1"
/db_xref="GI:14347200"
/translation="MPSASNTWENATLPSPFKDLPSPTNLLPURTSSNWT
CSLQTHFPKQYQPTSTSTATTPTPIKTTITTPPTNPLSDTNQPLPKW
NFLQAAATADLDVETALVSHERKHPKPKTADPRVQIAGNFAVPPEHAADQGLPVGK
IPKIDGVYVRNGANPLVEPVAGHFFDGDGMVAVKFTNGAASVACFTORLSOE
KSLGRPVPKAIGELHGHSGIARLLLFVARGLFGLVDSQCGMVGANGLVFNHLLA
MSEDDLPHYVRITNGDLTTVGRDFNGQLNSTMIARPKLPDVGDLHALSYDVLQKP
YLKTRFSPDGKSPDVEIPLKEPTMMHDFAITENFVVVPOQVVKLTMITGSPV
VYDKNTSRFGILHKNADNARMWDAPDCFCFHLNMAWEEPEETEEVYVVGSCMTPA
DSIENECESLSVLSIRLNRTKSTRRIISDAEQVNLNLEAGVNNRKLGRKTFPA
YLALAEWPVKVSGFAKVDLLSGEVKMYGKEFGPELFLPNGKEDDGVILAFVHD
EKWKESELOIYNAONLKEASIKLPSRPVYFGHTFIHSLDLRKAQ"
BASE COUNT 447 a 595 c 442 g 355 t
ORIGIN

Query Match 44.4%; Score 798.4; DB 6; Length 1839;
Best Local Similarity 70.4%; Pred. No. 1.6e-226;
Matches 1098; Conservative 0; Mismatches 456; Indels 6; Gaps 2;

Qy 233 ccaagaataccaactaacaagatgaatttgcagagagcggcggcagcgttgg 292
Db 281 CCAACCAACCATTAACCTCAAAAAATGGAATCTTCTCCAGAAAGCCGTGCGACGGCTTGG 340
Qy 293 acgcggcggaggggttcccttgcagccacgagaagctacacccgcttccctaaaaacggctg 352
Db 341 ACCTGGTGAACCGCGCTCGTCTCGCACGAGCGCAAAACACCCGCTCCCCAAAACGGCG 400
Qy 353 atcctagtggtcagatcgccggaatttgcctcgggtgaatgaacagccgctccggcgta 412
Db 401 ACCCAGAGGTCCAAATCCCGGGAACCTTCGCGCCCGGTGCCGAGGATGCGCCGGATCAAG 460
Qy 413 atcttcccggttggtcggaataacttcccgatccatcaaggagtgatgtgcgcaacggag 472
Db 461 GACTCCCGGTGTCGGAATAATCCCAAAATGATTCAGCGCGGTGTCGTGCGCAACGGTG 520
Qy 473 ctataccacttcaagcggcgtgacaggtcaccaatttctcgacggagacggtatcggttc 532
Db 521 CCAATTCGCTCTACGAGCTGTGGCGGGCACCACTTCTTCACGCGCGACGCGCATGGTCC 580
Qy 533 acgcggtcaaatcgaaacaggttcagctagctacgcttgcgggttactcagactaac 592
Db 581 ACGCCGTGAAGTTACGAACCGCGCGCAGCTACGCTTCGCCGCTTCACCGAGACGCGAC 640
```

Qy 593 ggtttggttcaggaaacgtcaattgggtcgaccgggttttccccaaagccatcggtgagcttc 652  
Db 641 GTCCTCGCAGAGAGAAATCTTAGCGCGCCCGGTGTTCGGAAGCCATCGGGAGCTCC 700  
Qy 653 acggccacacccggattatgccegaactcatgtattctacgcgcagagctcgacgcggtatag 712  
Db 701 ACGGCCACTCGCGCATCGCGCGCTCCTCTCTTCTACGCGCGGTCTCTTCGGGGCTCG 760  
Qy 713 tggaccgggacacaggaacccggtgtagtaacgcgcgggttttgggtctatttcaatggcgggt 772  
Db 761 TTGATGGGTCCAGGGCATGGCGGTGGGAACGCGCGTCTCGTCTACTTCAACACCAACC 820  
Qy 773 tattggctatgctcgagagatattacattaccaggttcagatcaactcccaatggagatt 832  
Db 821 TCTTGCCATGTCCGAAGACGATTTACCTTACCAGTGAAGATCACCTTAACGGCGACT 880  
Qy 833 taaaaacggttggtcggttcgatttggatggacaattagaaatccacaatgatgccccacc 892  
Db 881 TAAACACCGTTGGCGTTAGCACTTCAACGGGCAGCTCAACTCAACAATGATCGGCCACC 940  
Qy 893 cgaagtcgacccggaatcccggtgaactcttgcgttaagtacagactcggttcaaaagc 952  
Db 941 CGAAACTGACCCCTCGACGCGGACCTCCACGCGCTCAGCTACGACGTCAATTCAGAAGC 1000  
Qy 953 ctactctaaataactctcgattctcaccgacggaactaaataaccgacgctcgagattc 1012  
Db 1001 CTTACTCTAAGTACTTCGTTTCTCCCGCGGGCGTCAAGTCCCCCGACGTGGAATCC 1060  
Qy 1013 agcttgatcagccaacgatgatgcacgatttcgcgattacagagaacttcgctcgctac 1072  
Db 1061 CCCTGAAGAGCCACCATGATGCAGATTTCGCCATTAACGGAGAAATTCGTGCTGCTCC 1120  
Qy 1073 ctgaccagcaagtcgttttcaagtcgccggagatgatccgggtgggtctccgggtggtt 1132  
Db 1121 CCGACCAAGAGGTGTTCTTCAAACCTAACGGAGATGATCACCGCGGGTCCCCCGTGGTCT 1180  
Qy 1133 acgacaagaacaggtcgcaagattcgggattttagacaataaccgcaagatcatcga 1192  
Db 1181 ACGACAAGAACAAAACCTCACCGTTTGGGATTTCTGCACAAGAAATCGAAGGACGGAATG 1240  
Qy 1193 acatttaagtggattgatgctccagatgctctgtcttccatctctggaacgctttgggaag 1252  
Db 1241 CGATCGGTGGATGACGCGCGCGGATGTTTCTGCTTCCACCTCTGGAACCGTGGGAGG 1300  
Qy 1253 agccgaaacagatgaagtctcgtagatgggtctctgtagactccacagactcaattt 1312  
Db 1301 AGCCCGAAACCGAGGAGGTGTGGTGTATGGGTCTCTGATGACCCCTCGCGACTCCATT 1360  
Qy 1313 tcaacgagctgacagagaatctcaagagtgtcctgtctgaaatccgctgaatctcaaaa 1372  
Db 1361 TCACGCAATGCGAGGAGAGTTTGAAGAGCGTCTGTTCAGAGTAAGGCTGAACCTTGAGGA 1420  
Qy 1373 csggtgaatcaactcgcgctcgatcatctccaacgaagatcaacaagtcaactcgaag 1432  
Db 1421 CGGCAAGTCCACTCGGCGCCCATTAATCTCGACGCGG---AACAAGTGAACCTGGAG 1477  
Qy 1433 caggatggtcaacagaaacatgctcgccgttaaaccaaatctgcttacttgggttttag 1492  
Db 1478 CCGCATGTTGAACAGAAACAAGCTCGGAAGGAAGACCCAGTTCGCGTATCTGGCTCTCG 1537  
Qy 1493 cggagcgtggcctaaagctcaggattcgctaaagttgatctcaactcaggagaagtta 1552  
Db 1538 CGGAGCCCTGGGCCCCAAAGTCTCGGCTTTGCGAAAGTTGATTTGCTGAGTGGGGAAGTGA 1597  
Qy 1553 agaaacatctttacggcgataaccgttacggagagagacgtctgtttctccccggagag 1612  
Db 1598 AGAAGTACATGTATGGAGAAGAGAGTTCGGTGGGAGCCTCTGTCTTCTCC---AAGC 1654  
Qy 1613 gaggaggaagacaagagatacatctctgttttcggtcaagacagagaagacatggaat 1672  
Db 1655 GCCAAAAGAAAGACATGGGTATTTCTGGCATTTCTGTCACGACGAGAAAGAAATGGAAT 1714  
Qy 1673 cggagttacagatagtttaacgcggttagcttagaggttgaagcaacggttaaaccttcgt 1732

Db 1715 CCGAGCTGCAGATTGTGAATGCCCAAATTTAAAGCTCGAAGCTTCATCAAACTCCCT 1774  
Qy 1733 caagggttcgtagcggatttcacggtacattcatcgagccgatgatttggcgaagcagg 1792  
Db 1775 CTCGTGTTCCCTACGGTTTTTCATGGAACCTTTCATTTCATTCACAGGATTTCAGGAACAAG 1834  
RESULT 10  
AB030293 Vigna unguiculata CPRD65 mRNA for neoxanthin cleavage enzyme,  
LOCUS complete cds.  
DEFINITION  
AB030293 AB030293.1 GI:9857289  
VERSION neoxanthin cleavage enzyme.  
KEYWORDS Vigna unguiculata cDNA to mRNA.  
SOURCE Vigna unguiculata  
ORGANISM Vigna unguiculata  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Papilionoideae; Phaseoleae;  
Vigna.  
REFERENCE 1 (sites)  
AUTHORS Iuchi,S., Kobayashi,M., Yamauchi-Shinozaki,K. and Shinozaki,K.  
TITLE A stress-inducible gene for 9-cis-epoxycarotenoid dioxygenase  
involved in abscisic acid biosynthesis under water stress in  
drought-tolerant cowpea  
JOURNAL Plant Physiol. 123 (2), 553-562 (2000)  
MEDLINE 20317197  
REFERENCE 2 (bases 1 to 2432)  
AUTHORS Iuchi,S.  
TITLE Direct Submission  
JOURNAL Submitted (16-JUL-1999) Satoshi Iuchi, RIKEN, Plant Mol. Bio.;  
3-1-1 Koyuadai, Tsukuba 305-0074, Japan  
(E-mail:iuchi@tc.riken.go.jp, Tel:81-298-36-4359)  
FEATURES  
source 1..2432  
/organism="Vigna unguiculata"  
/db\_xref="taxon:3917"  
126..1964  
/genes="CPRD65"  
126..1964  
/genes="CPRD65"  
/codon\_start=1  
/product="neoxanthin cleavage enzyme"  
/protein\_id="BAB11932.1"  
/db\_xref="GI:9857290"  
/translation="MPSSASNTWFNATLSPFPKDLPTSSPTNLLPLRLKTSSTNIT  
CSLQTLHFPKQVOTSTSTSTATTTPTKTTTTPPTNPLSDNGLPVQK  
NFLKAAATALDAVETALVSHKHPKLPKADPRVOIAGNFAPVPEHAADQGLPVVQK  
IPKIDGVYVRNGANFLYPERVAGHHFFDGDGMVHAKFTNGAASACRFTETRLSQE  
KSIGRPVFKAIQELHGHGIRLLFYARGLFGLVDGSGMGVANAGLYVFNHLLIA  
MSDDLFPYHVRITPNDLTITVGRYDFNGQNSTMIARPKLDLPDVGDLHALUSYDIQKP  
YLKFRFSPGVKSPDVEIPLKEPTNHDFAITENFVVDQDQVFKLTMTIGSSPV  
VYDNKNTSREFGLHKNAKDANRWIDAPDFCFHLNWAEEPEVEVVGSMTPA  
DSIPLNCEESLSKVLSEIRLNTGTGSTRPPIISDAEONVLEAGWNRNKLGRKTPA  
YLALAEFPKRVSGFAKVDLLSGEVKKYMGEEKFGGEPLPLPNQKEDDGYILAFVHD  
EKEWSELQIVNAQNLKLEASIKLGRVPYGFHGTFTHSKDLRKA"  
BASE COUNT 654 a 711 c 533 g 534 t  
ORIGIN  
Query Match 44.4%; Score 798.4; DB 8; Length 2432;  
Best Local Similarity 70.4%; Pred. No. 1.7e-226;  
Matches 1098; Conservative 0; Mismatches 456; Indels 6; Gaps 2;  
Qy 233 ccaagaatccaacactaaacagatgaattgttccagagagcgcgcgcagcgttg 292  
Db 406 CCAACCAACCAATACCTCAAAAATGGAACCTTCTCCAGAAAGCCGCTGCCACGSCCTTGG 465  
Qy 293 acgcggcgagggtttccttgcagccacagagaagctacacccgccttctctaaacgcgtg 352  
Db 466 ACCTGGTCGAACCGGCTCGTCTCGCACGACGCAAAACACCCCGCTCCCAAAACGGCGG 525

Qy 353 atcctagtgttcagatgcccggaattttgtctccggtgaatgaacagcccgccgagcga 412  
Db 526 ACCGAGGGTCCAAATCGCGGGAATCTTCGCGCCGGTCCCGAGCATGCGCGGATCAAG 585  
Qy 413 atcttccggtggtcggaaacttcccgattccatcaaaaggaagtgtatgacgaacgag 472  
Db 586 GACTCCCGGTGTGCGAAAAATCCCCAAATGATTCAGCGCGGTGTACGTGCGCAACGGTG 645  
Qy 473 cttaacccacttcaagacgggtgacaggtcaccacttcttcgacggagacggtatggttc 532  
Db 646 CCAATCCGCTACGAGCTGTGGCGGGCACCATTCTTCGACGGCGGCGATGGTCC 705  
Qy 533 acgcggtcaaatcgaacacgggttcagtagtactacgcttgcgggtttactcagactaac 592  
Db 706 AGCGCGTGAAGTTCAGAACCGCGCGCGCAGCTACGCGCTTCACCGAGACGCAGC 765  
Qy 593 ggttggctcagaaacgtcaattgggtcgacgggttttcccaaaacgcatcggtgagcttc 652  
Db 766 GTCTCTCGAGGAGAAATCTCTAGCGCCCGGTGTTCGGAAGGCCATCGGGGAGCTCC 825  
Qy 653 acggccacacgggtattgcccgaactcatgtatttcaagccagagctgagcgggtatag 712  
Db 826 ACGGCCACTCCGGCATCGCGGCTCCTCCTCTTCTACGGCGCGGTCTCTTCGGGCTCG 885  
Qy 713 tcgacccggcaacacgggtgtagtaacgcgggtttgggtctatttcaatggccgggt 772  
Db 886 TTGATGGGTCCAGGGCATGGCGGTGGCGAACCGCGGTCTCGTCTACTTCAACACCAACC 945  
Qy 773 tattggtatgtcggagagatgattacacttaccgaagtcagactcctcccaatggagatt 832  
Db 946 TCTTGGCATGTCCGAAAGCAGATTTACCTACCGTGAAGATCACCCCTAACCGCGACT 1005  
Qy 833 taaaaacggttgggtcggttcgatttggatgacaaattgaaatcccaaatgattgcccacc 892  
Db 1006 TAAACACCGTTGGCGGTACGACTTCAACGGCGACTCAACTCAACATGATCGCCACCC 1065  
Qy 893 cgaagtcgacccgggaatcccggtgaactcttcgcttttaagctacgacgtcggttccaaagc 952  
Db 1066 CGAACTGGACCCCGTCAGCGCGACCTCCACGGCTCAGCTACGACGATTCAGAAAGC 1125  
Qy 953 cttaactaaaatacttcgattctcagcgacgggaactaaatcacggagctcgagattc 1012  
Db 1126 CTTACTCAAGTACTTCGTTCTCCCCCGACGGGTCAAGTCCCCCGACGTGGAATTC 1185  
Qy 1013 agcttgatcagcgaacgatgatgcacgatttcgcatcacagagaactcgtcgctgac 1072  
Db 1186 CCCTGAAGGAGCCCAACATGATGACGATTTCCCATTAACGGAGAAATTCGTCGTCGTC 1245  
Qy 1073 ctgaccagagtcgttttcaagctcgggagatgatccgcggtgggtctccggtgggtt 1132  
Db 1246 CCGACGACGAGTGTCTTCAAACCTAAACGAGATGATCACCGCGGTCCCCCGGTGCT 1305  
Qy 1133 acgacaagaacaggtcgaagattcgggattttagacaaataacgcgaagattcatcga 1192  
Db 1306 ACGACAAGAACAAAACCTCACCGTTTGGGATTTCTGCACAGAAGATCGAAGGACGCGAATG 1365  
Qy 1193 acattaaagtattgatgctccagattctctctgcttcattctcgttggaacgcttgggaag 1252  
Db 1366 CGATCGGTGATTCAGCGCGCGGATTTGTTCTGTTTCCACTCTGGAACCGGTGGAGG 1425  
Qy 1253 agccgaacacagatgaagtcgctgtagagggtcctgtagtactcaccagactcaattt 1312  
Db 1426 AGCCCGAAACGAGGAGGTGTGGTGTGTTGGTCTCTGATGATGATGATGATGATGATGAT 1485  
Qy 1313 tcaacgagctgaacagaaatctcaagagtgctcctctcgtgaatccgcctgaatcctcaaaa 1372  
Db 1486 TCAACGAATGCGAGGAGAGTTTGAAGAGCGGTCTGTCAGAGTAAGAGGTGAATTCAGGA 1545  
Qy 1373 cgggtgaatcaactcgcgcgtccgatacctcgaacgaagatcaacaagtcgaactcgaag 1432  
Db 1546 CCGGAAGTTCACCTCGGCGCCCCATTATCTCCGACGCCG ---AACAGTGAACCTGGGAG 1602

Qy 1433 cagggatggtcaacagaaacatgctcgccgtataaaacaaatctgcttacttgggtttag 1492  
Db 1603 CCGCATGCTGAACAGAAACAAGCTCGAAGAGAACCCAGTTCGCTATCTGGCTCG 1662  
Qy 1493 cgaagcgtggcctaaagtctcaggaattcgttaagttgatctcaactactggaagatta 1552  
Db 1663 CGGAGCCCTGGCCCAAGTCTCGGCTTTGCGAAAGTTGATTTGCTGTGTTGGGAGTGA 1722  
Qy 1553 agaaacatctttacggcgataaacggttacggaggagagcctctgtttctcccgagaaag 1612  
Db 1723 AGAAGTACATGATGAGAGAGAGATTCGTTGGTGGGAGCCTCTGTTCTTCCC ---AAGC 1779  
Qy 1613 gagagaggaagacgaagatacatcctctgttttcgacacgagaagacatggaat 1672  
Db 1780 GCCAAAAGAGACGATGGGTATATTTCTGGCATTCGTCCACGACGAGAAAGAAATGGAAT 1839  
Qy 1673 cggagttacagatagtttaacggcttagcttagaggttgaagcaacggttaaacctccgt 1732  
Db 1840 CCGAGCTCAGATTGTGAATGCCCAAAATTTAAAGCTCGAAGCTTCATCAAACTCCCT 1899  
Qy 1733 caagggtccgtacggatttcacggttacattcatcggagcggatgatttggcgaagcagg 1792  
Db 1900 CTCGTGTTCCCTACGTTTTTCATGGAACTTTTCATTCATTCACAGGATTTGAGAAACAAG 1959

## RESULT 11

AF190462 2398 bp mRNA linear PLN 19-JAN-2000  
LOCUS Phaseolus vulgaris 9-cis-epoxycarotenoid dioxygenase (NCED1) mRNA,  
complete cds.  
ACCESSION AF190462  
VERSION AF190462.1 GI:6715256  
KEYWORDS Phaseolus vulgaris.  
SOURCE Phaseolus vulgaris.  
ORGANISM Phaseolus vulgaris  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Phaseolus.  
REFERENCE 1 (bases 1 to 2398)  
AUTHORS Qin, X. and Zeevaert, J. A.  
TITLE The 9-cis-epoxycarotenoid cleavage reaction is the key regulatory  
step of abscisic acid biosynthesis in water-stressed bean  
Proc. Natl. Acad. Sci. U.S.A. 96 (26), 15354-15361 (1999)  
JOURNAL 20079657  
MEDLINE 10611388  
PUBMED  
REFERENCE 2 (bases 1 to 2398)  
AUTHORS Qin, X. and Zeevaert, J. A. D.  
TITLE Direct Submission  
JOURNAL Submitted (27-SEP-1999) MSU-DOE Plant Research Laboratory, Michigan  
State University, Wilson Str., East Lansing, MI 48824, USA  
Location/Qualifiers  
FEATURES  
source  
1..2398  
/organism="Phaseolus vulgaris"  
/cultivar="Top Crop"  
/db\_xref="taxon:3885"  
1..2398  
/gene="NCED1"  
124..1971  
/gene="NCED1"  
/function="catalyzes the oxidative cleavage of the C-11,12  
double bond of 9-cis-epoxycarotenoids (C40) to yield  
xanthoxin (C15) and C25-apo-aldehydes"  
/codon\_start=1  
/product="9-cis-epoxycarotenoid dioxygenase"  
/protein\_id="AAF26356.1"  
/db\_xref="GI:6715257"

## gene

## CDS

QKPYLYFRFSADGVKSPDVEIPLKEPTMMHDAIFAITNFVVVDQVVVFLTBMITGG  
SPVVDNKTSSRFILDKNAKANDAMRWIDAPECFPHLANAMEEPETDRIIVIGSCM  
TPADSIFNECDLSKLSVLSEIRNLRTGKSTRPIISDAEQVNLGAVNKLGRKT  
QFAYLALAEWPVKSGFAKVDLFSGEVQKMYGEEKFGEPLFLPNGBEEDGTLILAF  
VHDEKEWKSLEQIVNAQNLKLEASIKLPSRPVPGFHGTFIHSDKLRQKA"  
BASE COUNT 627 a 687 c 516 g 568 t  
ORIGIN

Query Match		44.0%;	Score	791.6;	DB	8;	Length	2398;
Best Local Similarity		70.0%;	Pred. No.	1.8e-224;				
Matches		1096;	Conservative	0;	Mismatches	464;	Indels	6;
						Gaps		2;
Qy	227	ccaaagccaaagatcccaactaaacagatgaattgttccagagagcgcgcgagcag	286					
Db	407	CGACACCAACCAACATTTCTCNAAAATGAACTTTCTGCAGAAACCCGCTCGCAG	466					
Qy	287	cgttggacggcgaggggttccctgtcagccacgagaagctacacccggttctcctaaaa	346					
Db	467	GCTTGGATATGTTGGAGACGGCGTGGTTTCGCACGAGCAACACCCCGCTCCCAAAA	526					
Qy	347	cgggtgacctagtggttcagatcgccggaattttgtctccgggtgaatgaacagcccgctc	406					
Db	527	CGCGGACCCGAAGTCCAAATCGCGCGGAATTCGCGCGCGTGCAGGACGACGCGCGG	586					
Qy	407	ggcgtaattctccggtggtcgaaaaactcccgattccatcaaaaggagtgatgtgcga	466					
Db	597	ATCAAGCCCTTCGGGTGTCGGAATAATTCCTCAATGATTCAGCGCGGTGACGTGCGCA	646					
Qy	467	acggagctaacccacttcacgagcgggtgacaggtcacactttctcgacggagacggta	526					
Db	647	ACGGCGCAATCCGCTCTACGAGCCCGTCGCGGGCACCACTTCTTCGACGGCGACGCA	706					
Qy	527	tgggttaacgcgctcaaatccgaacacggttcagctagtagtcagcttgcgggtttactcaga	586					
Db	707	TGTTCCACGCGCTGAAGTTCAGAACGCGCTGCGCAGTACGCGTGCCTTCACCGAAA	766					
Qy	597	ctaaccggttggttcagaaagctcaattgggtcgacaggttttcccaaaagccatcggtg	646					
Db	767	CACAGCGTCTCGGCAAGAGAAATCTTTAGCGCGCCCGGTGTTCCCGAAGCTATTCGCG	826					
Qy	647	agcttcacggccacacgggtattgcccgaactcatgctatttcaagccagagctgagcgg	706					
Db	827	AATCCACGGCCACTCCGGCATAGTTCGCTCCCTCTCTTCTACGCTCGACGCTCTTCC	886					
Qy	707	gtatagtcgacccggcagacacgggaaccggtgtagctaaacgcgggttggcttatctaatg	766					
Db	887	AGCTCGTTGATGGGTCCACGCGCATGGGGGTGGCGAAGCGCGGTCTCGTCTATTTCACA	946					
Qy	767	gcccgttaattggctatgtcggagagatgattacattaccattaccaagttcagatcaactccaatg	826					
Db	947	ACCACCTTTGGCCATGTCCGAAGACGATTTACCATACACGAGATACCTCTTAACG	1006					
Qy	827	gagatttaaaacgggttggtcggttcgattttgtatggacaattagaatcccaatgatgg	886					
Db	1007	CGACTTAACACCGTTGCGCGTTACGACTTCAAGCGCCAGCTCAACTCCCAATGATCG	1066					
Qy	887	cccacccgaaagtgcgacccggaatcccggtgaaactcttcgctttaagctacagctcggtt	946					
Db	1067	CCACCCGAAAATCTGACCCCGTCAACGCGGACCTCCACGCGCTCAGCTACGAGCTCGTTC	1126					
Qy	947	caaagcttaactaaatcttcgattcttcacggacgggaactaaatcacggagcgtcg	1006					
Db	1127	AGAGGCTTACTCTCAAGTACTTTCGCTTTCTCCGTGACGGCGTAAAGTCCCCCGACGTCG	1186					
Qy	1007	agattcagcttgatcagccacgatgatgcagatttcgcatgatttacagagaacttcgctg	1066					
Db	1187	AAATCCCTTGAAGAGCCACCATGATGCAGATTTTCGCAATAACGAGAGATTTTCGTCG	1246					
Qy	1067	tcgtaacctgaccagaagctcggttttcaagctgcccggagatgatccggtgggtctccgg	1126					
Db	1247	TCGTCCCCGACGACGAGGTGTTCTTCAAACTAACAGAGATGATCACCGCGGCTCCCCCG	1306					

Qy	1127	tggtttacgacaagaacaggtcgcaagattcgggatttttagacaaataacgccgaagatt	1186
Db	1307	TGGTCTACGACACAGAACAAACCTCACGGTTTGGGATTTTGGATAAGAAATCGGAGGACG	1366
Qy	1187	catcgaaacattaaatggtgattgagctccagatcgcttctctgcttccatctcttggacgctt	1246
Db	1367	CGAATGCGATGCGGTGGATCGACGCGCGGAGTGTCTTCTGCTTCCATCTCTTGGAACGGT	1426
Qy	1247	gggaagacagaaacagatgaagtcgctgtagaggtcctgtatgactacccacagact	1306
Db	1427	GGGAGGAGCCTGAAACCCAGCAGATTTGTGGTATTTGGGTCTGTCATGACCCCTCGGACT	1486
Qy	1307	caattttcaacgagctctgacgagaaatctcaagagtgctcctgtctgaaatccgctgaaatc	1366
Db	1487	CCATTTTCAATGATGCGATGAGATTTTGAAGACGCTTGTCTGAGATAAGGCTGAAT	1546
Qy	1367	tcaaaacgggtgaatcaactcgcgctccgatcatctccaaagaagatcaacaagctcaacc	1426
Db	1547	TGAGGACGGGCAAGTCGACTCGCGGCCCATTTATCTCGACGCGG--AACAAAGTGAACC	1603
Qy	1427	tcgaagcaggatggtcaacagaaaaacatgctcgccgctaaaaacaaattcgcttacttgg	1486
Db	1604	TTGAAGCCGCGCATGCTGAACAGAAACAAAGCTAGGAAGAAACACTCAGTTCGCTTATCTGG	1663
Qy	1487	ctttagccgagcggctggcctaaagtctcaggattcgcgtaaagtgtgactcactactggag	1546
Db	1664	CTCTGGCGAGCCCTGCGCCCAAGTCTCGGGCTTTGCGAAAGTTGATTTGTTTCAGTGGGG	1723
Qy	1547	aagtaagaacatcttttacggcgataacccgttacggagagagacccctctgtttctcccg	1606
Db	1724	AAGTGCAGAGTACATGTATGGAGAGAAAGTTCCGTTGGGGAGCCTCTGTTTCTTCCC-	1782
Qy	1607	gagaagagagagagagagagagagatcacatcctctgtttcgtttcacgagagagagacat	1666
Db	1783	--AACGCGAGAGAGAGAGTGTATTTCTGGCATTTCTGTCACGACGACGAGAAAGT	1840
Qy	1667	ggaatcggaggttacagatagttaaacccgcttagcttagaggttgaagcaacggtttaaac	1726
Db	1841	GGAATTCGAGCTCGAGATTTGTGAATGCCCAAAATTTAAAGTCGAAGCTTCCATCAAC	1900
Qy	1727	tccgctcaaggggttcgtracggatttcacgggtacattcatcggagcggatgatttggcga	1786
Db	1901	TCCCTCTCGGTTCTTATGGTTTTCATGGCACTTTTATTCATTCTAAGGATTGAGGA	1960
Qy	1787	ageagg 1792	
Db	1961	AGCAAG 1966	

RESULT 12  
AC074176/c  
LOCUS Arabidopsis thaliana chromosome 1 BAC T2H7 genomic sequence,  
DEFINITION complete sequence.  
AC074176  
VERSION AC074176.5 GI:12321625  
KEYWORDS HTG,  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
1 (bases 1 to 72058)  
Lin, X., Kaul, S., Town, C.D., Benito, M.-I., Creasy, T.H., Haas, B.J.,  
Wu, D., Maiti, R., Ronning, C.M., Koo, H., Fujii, C.Y., Utterback, T.R.,  
Barnstead, M.E., Bowman, C.L., White, O., Nierman, W.C. and Fraser, C.M.  
Arabidopsis thaliana chromosome 1 BAC T2H7 genomic sequence  
Unpublished  
2 (bases 1 to 72058)  
Town, C.D. and Kaul, S.  
Direct Submission  
Submitted (15-JUL-2000) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA, cdtown@igr.org



CDS	complement(join(19380..19710,19799..20403)) /gene="T2H7.4" /codon_start=1 /product="unknown protein" /protein_id="AAGS0858.1" /db_xref="GI:12321632" /translation="MKLVTAADKIPYDLYDAQLVGASHGWLFSDRNRSALISDYLN PYASKSPKMIHLPPFFPMYSGOTVEVNCVMSPPDQDDHEDHVDVVGIFLGROLS LCRRPDLRWNLITPESNEVSELMYSKDKQRFLLAPGNVILCSWDLNFKEDKRPK FHELVHLNPNRSLWKQDLSYGRDHWVESPSGESFLVWKFCEYSPENTSNWCKSP IWMVFEETKGRKNRYYDDIGLCIFISKGEDFCVKASCPGLQPNISLSHGRLF AIIILTKRTMGCVEYPOGIRIPILPYLPLPPLPSP" repeat_region complement(21599..21620) /rpt_family="AT_rich" /rpt_family="AT_rich" complement(21745..21765) /rpt_family="AT_rich" complement(join(21856..22030,22534..22806,23042..23208, 23290..23520,24442..24837)) /gene="T2H7.5" complement(21856..24837) /gene="T2H7.5" /codon_start=1 /product="En/Spm-like transposon protein, putative" /protein_id="AAGS0859.1" /db_xref="GI:12321633" /translation="MSSSSAIKAAKALERIMDTRVRSYVGLFSSVTVDMSKNRF RKQCIVLQIARGVARALNIPEDAVEVKTILNSWTCVEIVHTVKTPMAAPPKRGMPD
repeat_region	
repeat_region	
mRNA	Query Match Best Local Similarity 41.7%; Score 750.8; DB 8; Length 72058; Matches 1075; Conservative 0; Mismatches 492; Indels 9; Gaps 2;
gene	
CDS	223 aagcccaagccaaagaatccaaactaaacagatgaattgttccagagagcgcggcg 282 Db 55440 AACCCCAATAATCCGAAAAAGACACTTCCCGTTGGAAACCTCTCCAAACGCGCGCCCTCC 55381 Qy 283 gcagcgttggacgcggcgagggttccctgtcagccacgagaagctacacccgcttctc 342 Db 55380 CCGCACTCGACTTCGCGAAACCGCATGTGTAGACGCGGAACGTTCTTAACCTCTCCCT 55321 Qy 343 aaaaacggctgactctagtggttcagatccgcggaaatttgcctcgggtgaatgaacagccc 402 Db 55320 AAAACAGTTGATCTCTCGGCATCAGATCTCCGTAACACTACGCTCCGGTACCGGAGCAATCC 55261 Qy 403 gtcggcgtaattctccggttggtcgaaacttccgattccatcaaaaggagtgtatgtg 462 Db 55260 GTTAATCATCTCTCCGTTGATGGAATAATCCCTGACTGCAITGACGGTGTATTACCTC 55201 Qy 463 cgaacggagctaacccactcagagccggtgacaggtcaccacttcttcgacggagac 522 Db 55200 CGTAACGGCGCTAATCCACTCTTCGAGCCAGTTCTGTGTCATCACCTATTTCGACGGTGAC 55141 Qy 523 ggtatggttcacgcgctcaaatccgaacacgcggttcagtagtagctacgcttgcgggttact 582 Db 55140 GGTATGGTTACGCGCGTTAAAAATCACTAACGGAGACGCGAGTTACTCTGCGCGGTTTACG 55081 Qy 583 cagactaacgggttcttcaggaacgtcaattgggtcagccggttttccccaagccatc 642 Db 55080 GAAACCGAGATTGGTTTCAGAGANAAACAATCGGTTCTCCGATTTTCCCTAAAGCTATA 55021 Qy 643 ggtgagcttcacggccacacccggttatgtcccgactcatgtctattctacgccagagctgca 702 Db 55020 GGTGAGCTACATGGTCACTCTTGGATCGACGGTTGATGCTATTTTACGACGCTGTTTA 54961 Qy 703 gccggtatagtcgaacccggcacacgggaacccggttagtgaacgcggttttggtctatttc 762 Db 54960 TTCGGTTTATTAAATCACAANAAACCGGAGTTGCTAACCGCGGTTTGGTTTACTTTC 54901
CDS	763 aatggccgggttattggctatgtcggaggatgatttaacctaccaagttcagatcactccc 822 Db 54900 CACACCGGTTATTAGTATGTCTGAAGATGATCTACTCTTACCAGTTCGTGTCAGTGAC 54841 Qy 823 aatggagattataaaacccgttggtcgggttcgattttgatggacaattagaaatccacaatg 882 Db 54840 AATGGCGATTTAGAGACCATCGGAAGATTTCGATTTTCAGCGACAACCTAAGCTCCGCCATG 54781 Qy 883 attgccacaccgaaagtcgaccccggaatccgggtgaactcttcggttaagctaacgagctc 942 Db 54780 ATCCTCACCCGAAGATTGATCCGGTAACGAAGGAGCTATTTCGGTTGAGCTACGAGCTC 54721 Qy 943 gtttcaagccttacctaaaatacttccgatttccacgcgacggaactaaatccaccgac 1002 Db 54720 GTTAAGAAACCGTATTTCGAATACTTTAAATTTCTCGCGGAGGTGAGAATATCACCGGAC 54661 Qy 1003 gtcgagattcagcttgatcagcgaacacgatgatgcacgatttccgcgattacagagaatttc 1062 Db 54660 GTTGTGATTCGGGATCAACAAGTTGTGTTTAAAGCTCTCCGATATGTTTCTTGGGAAATCT 54541 Qy 1123 ccggtgggtttacgacaagaacaaggctcgcaagattcgggatttttagacaaatcgcgcgaa 1182 Db 54540 CCGGTTAAATACGACGGAGAGAAATTTCCGGTTTGGAAATTTTGCCTAGAAAACGCTAAA 54481 Qy 1183 gattcatcgaaacattaaagtgattgatgtccagattgtctctgcttccatctctcggaac 1242 Db 54480 GATCGCTCGGAGATGGTTGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 54421 Qy 1243 gtttgggaagagccagaaaacagatgaagtgcgtcgtgtagggctctgtatgactccaccca 1302 Db 54420 GCTTGGGAGTCCACGGACAGACGAGGTTGTAGTGTATCGGATCTTGCATGACTCCCGCG 54361 Qy 1303 gactcaattttcaacagagctgcagcagaatcgaagtgctcctgctgaaatccgcgctg 1362 Db 54360 GATTTCGATCTTCAACGAGTGGGATGAGCAGCTCAACAGTGTGTTTATCAGAGATCCGGTTA 54301 Qy 1363 aatctcaaaacccggtgaatcaactcgcgtccggtccggtatcttccaaacgaagatcaacaagtc 1422 Db 54300 AATCTCAAAACCGGGAATCCACACGACAGAACTATAATTTCCCGGTTCCGTT--CAGATG 54244 Qy 1423 aacctcgaagcaggatggtcacaacagaaaacatgctcggccgtaaaaaacaaatctcgcttac 1482 Db 54243 AATCTTGAAGCTGGTATGGTAAACCCGAAATCTTCTCGGAGAGAAAACCCGGTACGCATAC 54184 Qy 1483 ttggctttagccgagccgctggcctaaagtctcaggttcggttaagtgtgactcactact 1542 Db 54183 CTAGCCATAGCTGAACCGTGGCCTAAAGTCTCAGGATTCGCTAAAGTTGATCTTTCTTACC 54124 Qy 1543 ggagaagttaagaaacatctttacggcgataaccgtttacggaggagagcctctgtttcttc 1602 Db 54123 GGAGAAGTGAATAATCACTTTTACGCGGTAAAAATAACGGTGGTGAACCTTTTTTCTTA 54064 Qy 1603 cccgg-----agaaggaggaggaagaagaagatcacatcctctgttttcggttcagagac 1656 Db 54063 CCTAGAGGGTTAGAAATCTGACGGAGAGATGACGGTTACATATTATGTCGTTGTTGTTACGAC 54004 Qy 1657 gagaagacatggaaatcggagttacagatagtttaacccggttagcttagaggttgaagca 1716 Db 54003 GAGAGAGTTGGGAATCGGAGCTTCATATGTTTAAACCGCGTTAGCTTGGAACTAGAGCA 53944 Qy 1717 acggttaaaacttcggtcaagggttccgtagcgattttcacgggtacattcatcggagccgat 1776 Db 53943 ACCGTTAAATTCGCGTCAAGAGTACCGTATGCTGTTTCCACGCACTTTTCGTGAATTCGCGC 53884 Qy 1777 gatttggcgaagcagg 1792 Db 53883 GATATGTTAAACCCAGG 53868









```
/number=10
complement(7566..7702)
/gene="F28J12.20"
/number=11
8557..8838
/gene="F28J12.30"
/number=1
8557..13947
/gene="F28J12.30"
Join(8557..8838,8913..9050,9136..9198,9284..9373,
9450..9572,9668..9755,10089..10163,10261..10280,
10348..10368,12117..12725,12806..13321,13362..13622,
13693..13947)
/gene="F28J12.30"
/note="similarity to MucD - Pseudomonas aeruginosa,
PATCX:G1184684
/contains EST gb:T22208"
/codon_start=1
/product="putative protein"
/protein_id="CAA16717.1"
/db_xref="GI:2832642"
/db_xref="SPTREMBL:O49507"
/translation="MALASSKAFSSIFNTLSPINQSFVLACSGNSHVDVDRRRIM
IFGSSLATSSLLGNSOORLPWESAIALAEQKEELEEERENVALFKTSPSVY
IEALELPTSSGDLTDEBNGKIEGTGSGFWDKLGHIIVNIHVIKLAIDQGLQRC
IGNVYDAGTLTIFGWSGLGREIPSPNGKISIEAIOTDADINSNGSGLDVSYGHFI
GVNATFTRKSGMSGLIETDSVLNSVLTVMYERKKQIOIRNNSESNNRQXRRIS
HDKNRDELVYRILCPIDVVGVIKSGKVINAI RHNTAKI KVFQDLHGCSSORVIT
IYCSVKQOEIIGFTKSENEPLCCAQDALLAVYDAIVASDEENKTNDVDDNKECR
LLVPFSSSLIGKAGENIKRIRRTASVKVSKVDSPSHVCAMEYDNNVVISGSP
ESVKAOLFASAIWKYNPRENIPLDTSQDVPAAVSIVPSDLNSVYPQTFYSNOD
HILQOQAGPSYFNALSDPQGAETAANPVFPASSLPVTHFGGSSSEELVFKV
LCPILNMRVITKGGSTIKRIRREASGSCIEVNDRTKGDDECVIIIVATEILPCCLS
TPFVFMSPDDMKMAVEAVLLQEYINDEDAENKVMOLLVSVKVIKGVIGKSSVIN
EIRKTNANICISKGKDDLVVEVSEVSDPDLIOIIVLRREDVLDGKUSVATRKPP
ARTDNCFLSSSNAGYTLPSFMSMASTSGFHGIVGSPVVKSCILDFH"
8839..8912
/gene="F28J12.30"
/number=1
8913..9050
/gene="F28J12.30"
/number=2
9051..9135
/gene="F28J12.30"
/number=2
9136..9198
/gene="F28J12.30"
/number=3
9199..9283
/gene="F28J12.30"
/number=3
9284..9373
/gene="F28J12.30"
/number=4
9374..9449
/gene="F28J12.30"
/number=4
9450..9572
/gene="F28J12.30"

Query Match 41.0%; Score 738.2; DB 8; Length 110102;
Best Local Similarity 68.2%; Pred. NO. 3.2e-208;
Matches 1063; Conservative 0; Mismatches 478; Indels 18; Gaps 2;

Qy 244 aacactaacagatgaattgttcagagagcggcgagcgttgagcggcgag 303
Db 2180 AATCCTCTCCGCTAAACATCTTCAGAAAGCGCGCGATTGCGATCGACGCGCTGAG 2239

Qy 304 ggtttcctgtcagccagagaagctacaccgcttctctaaacggctgatcctagtgt 363
Db 2240 CGTGATTAACTCACAGACGACAGATTCTCACCATTCCAAAACCGCTGATCCAGTGT 2299
```

```
Qy 364 cagatcgccggaatttttgcctccggtgaatgaacagcccgctccggtgaattttccgggtg 423
Db 2300 CAGATTGCGGGAAATTAATTCCTCCGCTACCGGAATCTTCCTGTCGGCGAAACCTCACCGTC 2359

Qy 424 gtcggaataactcccgattccatcaagagtgatgctgcgaacggaggaactaacactt 483
Db 2360 GAAGGAACAATCCCTGACTGCTGACCGGTGTTATATCCGTAAACGGCGGAATCCGATG 2419

Qy 484 caccgagcgggtgacaggtcacactcttcgacggagacggttatggttcacgcccgtcaaa 543
Db 2420 TTTGAGCCAAACAGCTGGGACCACTTATTCGACGGAGACGGAATGTTTCACGCACTAAA 2479

Qy 544 ttcgaacacgggttagctagctagcttgcgggtttatctcagactaacgggtttgttcag 603
Db 2480 ATAACCAACGGTTACGTAGCTACGCATCGCGGTTTACAAAACCGAGAGATTGGTTTCA 2539

Qy 604 gaaagtcgaattgggtcgaccggttttcccaaacccatcggtgagcttcacggccacacc 663
Db 2540 GAAAACGATTGGGTGACACAGTTTTTCCGAAAGCAATCGGCGAGCTTCACCGTCACCTCG 2599

Qy 664 ggtattgcccgaactcatgctattctacgccagagctgcagcgggtatagtcgaccggca 723
Db 2600 GGAATCGCACGTTTTCATCTGTTTACGCACGTGGGCTTTGTTGTTCTGATCAACAACCAA 2659

Qy 724 caccgaaacgggtgtagctaacgcgggtttgtctctatttcaatggcgggttatgtggtatg 783
Db 2660 AACCGCGTCGGAGTAGCAAAACGCGGTTTGTGTTTACTTTAATAACCGGCTTTTAGCTATG 2719

Qy 784 tcgagagatgattacattaccgaagtccagatcactcccaatggagatttataaacgggtt 843
Db 2720 TCAGAAGACGATTTACCGGTACCAATTAATAAATTAACAAACGGCGGATCTCCAAACCGTT 2779

Qy 844 ggtcggttcgattttgatggacaattagaatcccaaatgattgccaccgcaaaagtcgac 903
Db 2780 GACGTTACGATTTTCGACGGTCAGTTAAATCCGCAATGATAGCTCACCCGAAACTGAC 2839

Qy 904 ccggaatccggtgaactcttcgctttaagctacgagctcggttcaaaagccttacctaaa 963
Db 2840 CCGGTTACGAAGAGGACTTTCACGCGTTAAGTACGACGTCGTTAAGAAACCTTACCTGAAA 2899

Qy 964 taactcgattctcacggagcgaactaaatacccgagcgtcgagattcagcttgatcag 1023
Db 2900 TACTTTCAGATTCTCGCGAGACGGCGTTAAATCGCGGAAATGGAGATCCCGCTCGAAACT 2959

Qy 1024 ccaacgatgacgacttttcgctgattacagagaacttcgctcgctgacctgaccagcaaa 1083
Db 2960 CCGACGATGATTCAGATTTTCGCTATAACGGAGAAATTTTGTGGTGATTCCTGATCAACAA 3019

Qy 1084 gtcgttttcaagctgcgggagatgatccgcgggtgggtctccgggtggtttacgacaagaac 1143
Db 3020 GTCGTGTTCAAAGCTCGGCGAGATGATTTTCGCTAAATCTCCGGTTGTTTTCGACGGAGAA 3079

Qy 1144 aaggtcgcaagattcgggatttttagaataaacgccaagattcattcgaaataaagtg 1203
Db 3080 AAGGTTTCCCGATTGGGGATAATGCCCAAGACCGCAGACAGAGCTTCTCAGATAATCTGG 3139

Qy 1204 attgatgctccagatgctctctgcttccatctctggaaacgcttgggaagagcgaagaaca 1263
Db 3140 GTGAATCTCCGGAGACGTTCTGTTTTCATCTCTGGAATGATGGGAATCGCCGGAGAGCG 3199

Qy 1264 gatgaagtgcgtgatagggctcctgtatgactccaccagactcaattttcaacagactct 1323
Db 3200 GAGGAGATTGTGGTGATCGGATCGTATGTCGCGCGGGGATTCATCTTCAACGAGAGA 3259

Qy 1324 gacgagaatctcaagatgctcctgtctgaaatccgctgaaatcctcaaaacgggtgaaca 1383
Db 3260 GACGAGAGCTTGAGAAGCGTTTGTTCGAGATCAGGATAAAACCTCAGAACACGTAATAACC 3319

Qy 1384 actcgcgctccgatcactctcaacgaagatcaacaagccaacctcgagcgggatgctc 1443
Db 3320 ACGGTCGTTCTGTTGTTGTTAACGAGGAT-----GTAATTTAGAGATTGGTATGTT 3373

Qy 1444 aacagaaacatgctcgccggtaaaaaaccaaatctgcttacttggttttagccgagccgtgg 1503
```

Db	3374	AACCGAACCGGTTAGGAAGAAACCCGGTTCGGGTTTGGCTATTGCTTATCCITGG	3433
Qy	1504	cctaaagtctcagattcgcgttaagtgtgatctcactactgagagaagtttaagaacatctt	1563
Db	3434	CCAAAAGTTTCGGGTTTCGCTAAAGTCGATCTTTGCACCGGTGAGATGAAAAATATATT	3493
Qy	1564	tacggcataaccgttacggaggagagcctctgtttctcccg------agaa	1611
Db	3494	TACGCGGTGAGAAATATGCGCGGAACCGTTTCTTGCCTCCGCAACTCCGGTAACGGC	3553
Qy	1612	ggagagagggaagcgaagatacatctctgttcttcacgcagcagaagacatggaaa	1671
Db	3554	GAAGAAATGAAGATGACGGTTATATTTGTACGTTTCATGACGAGAAACAAGACA	3613
Qy	1672	tcggagttacagatagtttaacccgcttagcttagagttgaagcaacggttaaacctccg	1731
Db	3614	TCAGAGCTTCAGATTATTAACGCTGTTAATTTAAAGCTTGAAGCTACCGATTAAACTACCG	3673
Qy	1732	tcaaggttcggtacgatttcacggatcattcgcgagccgagcgtgatttggcgaagca	1790
Db	3674	TCTAGAGTACCCTATGGGTTTCATGCGCACATTTGTGGATTGCAATGAACCTCGTTGATCA	3732
RESULT	15		
LOCUS	ATCHRIV48	194143 bp	DNA linear
DEFINITION	Arabidopsis thaliana	194143 bp	DNA linear
ACCESSION	AL161548		PLN 16-MAR-2000
VERSION	AL161548.2	GI:7268604	contig fragment No. 48.
KEYWORDS			
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1 (bases 1 to 194143)		
AUTHORS	Hilbert,H., Braun,M., Holzer,E., Brandt,A., Duesterhoeft,A., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 38835; 28388 to 111084)		
AUTHORS	Murphy,G., Ridley,P., Hudson,S., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 194143)		
AUTHORS	EU Arabidopsis sequencing project.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de		
COMMENT	Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk		
FEATURES	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ this fragment has an overlap with ATCHRIV47 at the 5' end and an overlap with ATCHRIV49 at the 3' end.		
source	1. 194143		
organism	"Arabidopsis thaliana"		
variety	"Columbia"		
db_xref	"taxon:3702"		
chromosome	"4"		
6377..9293			
gene	/gene="AT4g18120"		
CDS	join(6377..6796,7122..7238,7347..7501,7578..8456,8543..8627,8716..8784,8871..9293)		
note	"similarity to Mei2-like protein, Arabidopsis thaliana, gb:D86122 contains EST gb:W43784, AI998063.1"		
codon_start	1		

/product	"putative protein"
/protein_id	"CAB78814.1"
/db_xref	"GI:7268605"
/translation	"MASGLDHFSGGIGNMLDDGSDHPIGNMLPDDDEELFSGLMDDLN LSSLPATLDIEDYDLFGSGGLELTDYPDYSLNKGFSGMGFADSNVMPQNIQFQNGVSIAGEHPYGEHPRTILFVRNINSNVEDSELQALFEDNPSEKDVNQGTLVVFNLAPEVSNROLENI FGVIKEIQIRTPNKRHHKFVEFFDVRSAADAAKALNRTEIAGKRI KLEHSPGGARRNMLQMPLEQDDSYSLAHVESPLASSPIGNWRNSPIDHPLOSF SKSPIGNLSPTKNIRYPESMKTASVNDQEGRRSHDLHLFSSSSYNNASHKASTF QQQPSGVSFSSFGSLNSHPSHVETLSGSBFLWGPSSSAWVPVNPFSNRENHRFPYSA QNGLHQLHHIGSAPGFPFRSPETSMGSVAFRGASGNNNAQNRLETSSPNFQMLS APRRSOLFTGNSYLPAAATMWSIDDPLEDGNSQFDSNGQADIKIQFOLDLSKIMR GEDPRTLMKNIPNKTTRNMLIAAIDEKNSGTDFLYLPIDFNKCNVCYAFINMVS PETIAYEAFNGKWKDKENSEKVASIARVIOGKAALIAHFONSSIMNEDRRCQPIV FDGSESKYPVNNLLQQTDLKMLIYLCVYLMNWLTHCLFETRSGSEFAARASNSTG FLFLSARTHIKADSQNVNHIASGLKRRKYQFLG"
6377..6796	
/gene	"AT4g18120"
/number	1
6797..7121	
/gene	"AT4g18120"
/number	1
7122..7238	
/gene	"AT4g18120"
/number	2
7239..7346	
/gene	"AT4g18120"
/number	2
7347..7501	
/gene	"AT4g18120"
/number	3
7502..7577	
/gene	"AT4g18120"
/number	3
7578..8456	
/gene	"AT4g18120"
/number	4
8457..8542	
/gene	"AT4g18120"
/number	4
8543..8627	
/gene	"AT4g18120"
/number	5
8628..8715	
/gene	"AT4g18120"
/number	5
8716..8784	
/gene	"AT4g18120"
/number	6
8785..8870	
/gene	"AT4g18120"
/number	6
8871..9293	
/gene	"AT4g18120"
/number	7
complement(9594..9837)	
/gene	"AT4g18130"
/number	1
complement(join(9594..9837,9944..10237,10323..11130,11238..13230))	
/gene	"AT4g18130"
9594..13230	
/gene	"AT4g18130"
/gene	"AT4g18130"
complement(join(9594..9837,9944..10237,10323..11130,11238..13230))	
/gene	"AT4g18130"
/note	"Contains ATP/GTP-binding site motif A (p-loop) AA634-641; Phytochrome chromophore attachment site AA317-326 contains EST gb:T21020, AA395462"
/codon_start	1
/product	"phytochrome E"
/protein_id	"CAB78815.1"



```

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111045 AAGGTTTCCCGATTGGGGATAATGCCCAAGGCGGACAGAAAGCTTCTCAGATAATCTGG 111104

Qy 1204 attgatgtccagattgtttcttgcattccatctctggaacgcttggaagagccagaaaca 1263
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111105 GTAAACTCTCCGGAGACGTTCTGTTTTCATCTCTGGAATGCATGGGAATCGCCGAGACG 111164

Qy 1264 gatgaagtcgctcgatagggctcctgtatgactccacagactcaattttcaacgagctc 1323
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111165 GAGGAGATTGGTGATCGGATCGTGATGTGCGCCGGCGGATTCAATCTTCAACGAGAGA 111224

Qy 1324 gacgagaatctcaagagtgctcctgtctgaaatccgcctgaaatctcaaaacccgtgaaatca 1383
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111225 CACGAGACGTTGAGNAGCGTTTGTTCGAGATCAGGATAAACCTCAGAACACGTAATAACC 111284

Qy 1384 actcgcctccgatactctccaacgaagatcaacaagtcacaacctcgaagcagggatggtc 1443
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111285 ACGCTCGTTCGTTGTTGTTTAAACGAGGAT-----GTAAATTTAGAGATTGGTATGGTT 111338

Qy 1444 aacagaaacatgctcggcgtgaaaccaaattcgcttacttggcttttagccgagccgtgg 1503
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111339 AACCGAAACCGGTTAGGAAGAAAAACCCGGTTCGGGTTTTGGGCTATTGCTTATCCTTGG 111398

Qy 1504 cctaaagtctcaggattcgctaaagtgtgatctcactactggagaagttaaagaacatctt 1563
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111399 CCAAAAGTTTCGGGTTTCGCTAAGGTCGATCTTTGCACCGGTGAGATGAAAAATATATT 111458

Qy 1564 tacggcgataaccgttacggaggagagcctctgtttctcccgg-----agaa 1611
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111459 TACGGCGGTGAGAAATATGGCGGCGAACCGTTTTCTTTCGCCGCGCACTCCGGTAACGCG 111518

Qy 1612 ggaaggaggaagaaaggaatacatcctctgtttcgttcacgacgagaagacatggaaa 1671
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111519 GAAGAAAAATGAAGATGACGGGTTATATATTTGTACAGTTCATGACGAAGAAACAAAGACA 111578

Qy 1672 tcggagttacagatagtttaacgcccgttagcttagaggttgagcaacggttaaacctccg 1731
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111579 TCAGAGCTTCAGATTATTAAACGCTGTTAATTTAAAGCTTGAAGCTACGATTAACTACCG 111638

Qy 1732 tcaagggttcgtagcaggtttcacgggtacattcatcgagccgatgatttggcgaagca 1790
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111639 TCTAGAGTACCGTATGGGTTTCATGGCACATTTGTGGATTGGAATGAACTCGTTGATCA 111697

```

Result	No.	Score	Query Match	Length	DB	ID	Description
	1	1800	100.0	1800	22	AA009396	Arabidopsis
	2	824.6	45.8	1818	22	AA009401	Lycopersicon
	3	798.4	44.4	1839	22	AA009399	Vigna unguiculata
	4	738.2	41.0	1752	22	AA009394	Arabidopsis
	5	606.6	33.7	1815	22	AA009400	Zea mays
	6	528.2	29.3	1734	22	AA009398	Arabidopsis
	7	229.8	12.8	443	21	AAC56678	Eucalyptus
c	8	204	11.3	492	22	AA009406	Rice
	9	188	10.4	372	21	AAC56695	Eucalyptus

PI Iuchi S, Kobayashi M, Shinozaki K;  
XX WPI; 2001-400081/43.  
DR P-PSDB; AAE04784.  
XX  
PT A DNA encoding a protein with a neoxanthin cleavage activity for  
PT producing transgenic plants with improved or decreased stress tolerance  
PT -  
XX  
XX Claim 3; Page 32-36; 101pp; English.  
XX  
XX The invention relates to neoxanthin cleavage enzymes and their  
CC corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key  
CC role in endogenous abscisic acid (ABA) biosynthesis under drought stress.  
CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a  
CC plant when expressed in a plant cell. The invention also relates to  
CC methods for increasing or decreasing stress tolerance in a plant by  
CC introducing the DNA into the plant, and a transgenic plant into which a  
CC neoxanthin cleavage enzyme is introduced. The improvement of stress  
CC tolerance in plants is useful, for example, in plant breeding. Neoxanthin  
CC cleavage enzyme genes are useful for producing transgenic plants. An arid  
CC land can be improved by growing transformant weed for several years and  
CC then removing the weed by specifically lowering stress tolerance in the  
CC weed by inducing an inducible promoter. The present cDNA sequence encodes  
CC Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCD3 protein.  
CC The AtNCD3 cDNA is obtained from an Arabidopsis plant-derived cDNA  
CC library using a cDNA of the cPRD5 (Cowpea Responsive to Dehydration)  
CC gene isolated from cowpea plant as a probe.  
XX  
XX Sequence 1800 BP; 458 A; 464 C; 439 G; 439 T; 0 other;

Query Match 100.0%; Score 1800; DB 22; Length 1800;  
Best Local Similarity 100.0%; Pred. NO. 0;  
Matches 1800; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 atggtctttccacggaacggctgcgggtttctggagatggcttggtggaatcatact 60  
Db 1 atggtctttccacggaacggctgcgggtttctggagatggcttggtggaatcatact 60  
Qy 61 cagcgcgcatattcgctcttcctcaagctccgacttgattgttagctctcttacctatg 120  
Db 61 cagcgcgcatattcgctcttcctcaagctccgacttgattgttagctctcttacctatg 120  
Qy 121 gccagtcgtgtcaacagtaagctcaatgtttcatctgcgtctcacactctccagctctt 180  
Db 121 gccagtcgtgtcaacagtaagctcaatgtttcatctgcgtctcacactctccagctctt 180  
Qy 181 catttccctaagcaatcatcaactctccgccattgtgttaagcccaagccaaagaa 240  
Db 181 catttccctaagcaatcatcaactctccgccattgtgttaagcccaagccaaagaa 240  
Qy 241 tccaaactaaacagatgaatttttccagagagcggcggcggcagcgttggacgcggcg 300  
Db 241 tccaaactaaacagatgaatttttccagagagcggcggcggcagcgttggacgcggcg 300  
Qy 301 gaggggttccttgtcagccacgagaagctacacccgttctcctaaacggctgacctagt 360  
Db 301 gaggggttccttgtcagccacgagaagctacacccgttctcctaaacggctgacctagt 360  
Qy 361 gttcagatccgcggaatttttgcgtccggtgaatgaacagccgctccggtgaatcttccg 420  
Db 361 gttcagatccgcggaatttttgcgtccggtgaatgaacagccgctccggtgaatcttccg 420  
Qy 421 gtggtcggaaaaacttcccgattccatcaaaaggatgtatgtcgcaacggagctaaacca 480  
Db 421 gtggtcggaaaaacttcccgattccatcaaaaggatgtatgtcgcaacggagctaaacca 480  
Qy 481 cttcacgagccggtgacaggttcacacttcttcacggagacggtatggttcacgcgctc 540  
Db 481 cttcacgagccggtgacaggttcacacttcttcacggagacggtatggttcacgcgctc 540  
Qy 541 aaattcgaacgggttcagctagctacgcttgcgggttttactcagactaacgggttggtt 600

Db 541 aaattcgaacgggttcagctagctacgcttgcgggtttactcagactaacgggttggtt 600  
Qy 601 caggaaacgtcaattgggtcgacgggttttcccaaaagccatcggtgagcttcacggccac 660  
Db 601 caggaaacgtcaattgggtcgacgggttttcccaaaagccatcggtgagcttcacggccac 660  
Qy 661 accggtattcccgactcatgtattctacgacagagctcagccggtatagtcgacccg 720  
Db 661 accggtattcccgactcatgtattctacgacagagctcagccggtatagtcgacccg 720  
Qy 721 gcacacggaacgggtgtagctaaacgggttttggtctatttcaatggccggttattggct 780  
Db 721 gcacacggaacgggtgtagctaaacgggttttggtctatttcaatggccggttattggct 780  
Qy 781 atgtcggagagatgatttacctttaccaagttcagatcactcccaatggagatttaaaacc 840  
Db 781 atgtcggagagatgatttacctttaccaagttcagatcactcccaatggagatttaaaacc 840  
Qy 841 gttggtcgggttcgatttttgagacaattagaatccaatgattgccaccggaagtc 900  
Db 841 gttggtcgggttcgatttttgagacaattagaatccaatgattgccaccggaagtc 900  
Qy 901 gaccgggaatccgggtgaactcttcgcttttaagctacgactcgctttcaaaagctttacct 960  
Db 901 gaccgggaatccgggtgaactcttcgcttttaagctacgactcgctttcaaaagctttacct 960  
Qy 961 aaatacttcgattctcaccggacgggaactaaatcaccgagctcgagattcagcttgat 1020  
Db 961 aaatacttcgattctcaccggacgggaactaaatcaccgagctcgagattcagcttgat 1020  
Qy 1021 cagccaacgagatgacgatttcgggattacagagaacttcgtcgtacctgaccag 1080  
Db 1021 cagccaacgagatgacgatttcgggattacagagaacttcgtcgtacctgaccag 1080  
Qy 1081 caagtgcgttttcaagctccggagatgatccgggtgggtctccgggtggttttacgacaag 1140  
Db 1081 caagtgcgttttcaagctccggagatgatccgggtgggtctccgggtggttttacgacaag 1140  
Qy 1141 aacaagtcgcaagattcgggatttttagacaataacgccgaagattcatcgaaatcaag 1200  
Db 1141 aacaagtcgcaagattcgggatttttagacaataacgccgaagattcatcgaaatcaag 1200  
Qy 1201 tggattgactgcagatttcgtctcgtctccactctcggaaacttcgggaagaccgaa 1260  
Db 1201 tggattgactgcagatttcgtctcgtctccactctcggaaacttcgggaagaccgaa 1260  
Qy 1261 acagatgaagtgcgtgtagagggtcctgtatgactccaccagactcaattttcaacgag 1320  
Db 1261 acagatgaagtgcgtgtagagggtcctgtatgactccaccagactcaattttcaacgag 1320  
Qy 1321 tctgacgaagaatctcaagagtgctcgtctgaaatccgctgaaatctcaaaacggcgaa 1380  
Db 1321 tctgacgaagaatctcaagagtgctcgtctgaaatccgctgaaatctcaaaacggcgaa 1380  
Qy 1381 tcaactccggtccgactcatctcacaagagatcaacaagtcaacctcgaagcgggatg 1440  
Db 1381 tcaactccggtccgactcatctcacaagagatcaacaagtcaacctcgaagcgggatg 1440  
Qy 1441 gtaacagagaacatgctcggcgttaaaaccaaatttcgcttacttgggttttagccgagccg 1500  
Db 1441 gtaacagagaacatgctcggcgttaaaaccaaatttcgcttacttgggttttagccgagccg 1500  
Qy 1501 tggcctaaagtctcaggatttcgtaaaagttgatctactactgagagaagtaagaacat 1560  
Db 1501 tggcctaaagtctcaggatttcgtaaaagttgatctactactgagagaagtaagaacat 1560  
Qy 1561 ctttacggcgaataacggttacggagagagacccctgttctccccggagagggggag 1620  
Db 1561 ctttacggcgaataacggttacggagagagacccctgttctccccggagagggggag 1620  
Qy 1621 gaagacgaaggaatacatcctctgttctcgttcacgacgagaagacatggaatcggagtta 1680

Db 1621 gaagacgaaggatacactctgttctgttcacgacgagaagacatggaaatcgagggtta 1680  
 Qy 1681 cagatagttaaacgcgttagcttagaggttgagcaacggttaaaacttcctgcaagggtt 1740  
 Db 1681 cagatagttaaacgcgttagcttagaggttgagcaacggttaaaacttcctgcaagggtt 1740  
 Qy 1741 cagatagttaaacgcgttagcttagaggttgagcaacggttaaaacttcctgcaagggtt 1800  
 Db 1741 cagatagttaaacgcgttagcttagaggttgagcaacggttaaaacttcctgcaagggtt 1800

## RESULT 2

AAD09401

ID AAD09401 standard; cDNA; 1818 BP.

XX AC AAD09401;

XX 10-SEP-2001 (first entry)

DE Lycopersicon esculentum neoxanthin cleavage enzyme, LeNCED1 cDNA.

XX Tomato; neoxanthin cleavage enzyme; LeNCED1; abscisic acid; ABA;  
 KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
 KW plant growth protectant; herbicide; ss.

XX Lycopersicon esculentum.

XX Key Location/Qualifiers  
 FT 1..1818  
 CDS /\*tag= a

FT /product= "Lycopersicon esculentum LeNCED1 protein"

XX EP116794-A2.

XX 18-JUL-2001.

XX 11-JAN-2001; 2001EP-0300218.

XX 13-JAN-2000; 2000JP-0010056.

XX 11-JAN-2001; 2001JP-0003476.

XX (RIKE ) RIKEN KK.

XX Iuchi S, Kobayaashi M, Shinozaki K;

XX WPI; 2001-400081/43.

XX P-PSDB; AAE04789.

XX A DNA encoding a protein with a neoxanthin cleavage activity for  
 PT producing transgenic plants with improved or decreased stress tolerance

XX -

XX Claim 3; Page 67-71; 101pp; English.

XX The invention relates to neoxanthin cleavage enzymes and their  
 CC corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key  
 CC role in endogenous abscisic acid (ABA) biosynthesis under drought stress.  
 CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a  
 CC plant when expressed in a plant cell. The invention also relates to  
 CC methods for increasing or decreasing stress tolerance in a plant by  
 CC introducing the DNA into the plant, and a transgenic plant into which a  
 CC neoxanthin cleavage enzyme is introduced. The improvement of stress  
 CC tolerance in plants is useful, for example in plant breeding. Neoxanthin  
 CC cleavage enzyme genes are useful for producing transgenic plants. An arid  
 CC land can be improved by growing transformant weed for several years and  
 CC then removing the weed by specifically lowering stress tolerance in the  
 CC weed by inducing an inducible promoter. The present cDNA sequence encodes  
 CC Lycopersicon esculentum neoxanthin cleavage enzyme, LeNCED1 protein  
 CC related to the invention.

XX Sequence 1818 BP; 569 A; 383 C; 369 G; 497 T; 0 other;

Query Match 45.8%; Score 824.6; DB 22; Length 1818;  
 Best Local Similarity 69.6%; Pred. NO. 2.1e-256;  
 Matches 1134; Conservative 0; Mismatches 489; Indels 6; Gaps 1;

Qy 170 cccagctcttcattcccttaagcaatcatcaaacctcccgcattgttgaagcca 229  
 Db 185 cttcaaatatcaaacaccaaagaataatacaatttcacacccaaacaaagaaacaaca 244  
 Qy 230 aagccaaagaatcccaactaaacagatgaatttgcagagagcggcggcagcgt 289  
 Db 245 actctcttcttccaaactccaagtgaatttagtcgagaagcagcagaaggcct 304  
 Qy 290 tggacgcggcgagggtttcttctgtcagccacgagaagctacaccgcttctctaaacgg 349  
 Db 305 tagatgctgtagaagtgctttaactaaacatgaactgaaacaccttgcgcgaaacag 364  
 Qy 350 ctgaccttagtctcagatcgccggaattttgctcccggtgaatgaacagccgctccggc 409  
 Db 365 ccgacccacgagtcagatttctggaattttgctcccggtaccggaaaaatccagctgtc 424  
 Qy 410 gtaatcttcgggtggtcggaacttccgattccatcaaaaggagtgtatgctgcgaacg 469  
 Db 425 aatctcttcgggtcaccggaaaaataaccaaatgttcaaggcgttacgttcgaaacg 484  
 Qy 470 gagctaacccattcacgagcgggtgacaggtcacacttcttcagcggagacggtatgg 529  
 Db 485 gagctaacctcttttgaaccaaccgcggacacatttcttcgacggcggcgtatgg 544  
 Qy 530 ttcacgcgctcaaatcgaaacggttcagtagctacgcttcggttctactcagacta 589  
 Db 545 ttcacgcgctcaaatcgaaacggttggctgagttacgcttgcggttctactcagacta 604  
 Qy 590 accggttctcaggaaacgtcaattgggtcgaccggttttccccaagccatcggtagc 649  
 Db 605 agagcgttctcaagaaaaagctttgggtcgccctgtttccctaaagcattggtgaat 664  
 Qy 650 ttcacggccacacgcgttatcccgactcatgctatttaacgccagagctgcagccgcta 709  
 Db 665 tacatggtcaactctggaattgcaaggctatgctgttttaacgctcggtggtcttcggac 724  
 Qy 710 tagtcgacccggcacacggaacgggtgagtaacgcgggtttggtctatttcaatggcc 769  
 Db 725 ttgttgatcacagtaaaaggaaactggtgtgcaaacgcgggttttagctatttcaataacc 784  
 Qy 770 ggttattggtctatgctggaggatgatttacccttaccaggttcagatcactcccaatcgag 829  
 Db 785 gattactgctatgctgaagatgattgcttaccatgataaggttaacacccaccggcg 844  
 Qy 830 atttaaaacgcgttgctcggttcgattttagtggaacattagatcccaaatgattgccc 889  
 Db 845 attcttaaacagaggtcgatttcgacggcagctaaaatccaccatgatagctc 904  
 Qy 890 accgaaagtgcacccggaatccggtgaactcttcgcttttaagctacgacgcttcaaa 949  
 Db 905 acccaaaagtgcacccagttcccggtgagctatttgccttagctacgattgattcaga 964  
 Qy 950 agccttaccttaaaatctccgattctcaccgacggaaactaaatcacccgacgtcgaga 1009  
 Db 965 agccatacctcaagctacttcagattttcaaaaaatgggaaaaatcaaatgatgtgaaa 1024  
 Qy 1010 ttcagcttgatcagccacgatgatgcacgatttcgcgattacagagaacttcgctcgtcg 1069  
 Db 1025 ttccagtggaaccccaaatgatgcatttgcgaaattactgagaacttcgctcgtcga 1084  
 Qy 1070 tacttgacagcaactcgttttcaagctgcggagatgatccgcggttggttccgggtgg 1129  
 Db 1085 ttccgtgacaaagtcgttttcaagatgctgaaatgatccgtggaggttccacgggtgg 1144  
 Qy 1130 tttacgacaaagaacaggtcgcaagatttcggaatttttagacaaaatcacgccgaagattcat 1189  
 Db 1145 tttacgacaaagaacaggttcccgatttgggtattctctggataagtcacgcgaagatgggt 1204  
 Qy 1190 cgaacattaagtggtgatgctccagattgtcttctgtcttccatctctggaacggttggg 1249

Db 1205 ctgatttgaaatgggtgaagtacctgattgtttctgtttccacctctggaatccttggg 1264  
Qy 1250 aagaaccgaagaacagatgaagtcgtcgtagaggtctctgtatgactcacacagactcaa 1309  
Db 1265 aagaagcgaagaacagatgaatcgttgaattggttcattgtagacaccacagactcca 1324  
Qy 1310 ttttcaacgagctgacagagaatccaagatgctcctctgaaatccgctgaatctca 1369  
Db 1325 ttttcaatgaatgtagaagggtcaagagtggttttatccgaaatccgtctcaatttga 1384  
Qy 1370 aaaccggtgaatcaactcgcgtcgcatctctcaacgaagatcaacaagtcaacctcg 1429  
Db 1385 aaacagggaatcaacaagaataatcataatcgaaaacccggatgaacaagtgaatttag 1444  
Qy 1430 aagcaggatggtcaacgaagaacatgctgcgcgttaaacacaaatctccttacttgggt 1489  
Db 1445 aagctggaatggtgaaccgaagaacactcgaaggaagaacagatgcttatttggcta 1504  
Qy 1490 tagcgaagcgtggtcctaagctcaggatctcgctaaagtgtgactcactactggagaag 1549  
Db 1505 tgcgtgaacctggccaaagtcttctgttttgcaaaagtataacctgtctcccggtgaag 1564  
Qy 1550 ttaagaacatctttacggcgataaccgttacggaggagagcctctgtttctcccggtgag 1609  
Db 1565 ttgaagaattcatttatggtgacaacaataatggtgggaacacctctttttaccagaag 1624  
Qy 1610 -----aaggaggagagaagaacagatatacatctctgtttcgttcacacagagaaga 1663  
Db 1625 accccaacagcaaggaaagacgagtggttatatttttgcgttcacgatgagaag 1684  
Qy 1664 catggaatcggagttacagatagtttaacgcgttagcttagaggttgaagaacaggtta 1723  
Db 1685 aatggaatcgaactgcaaatgtttaacgcaatgagtttgaagttgaggcaactgtga 1744  
Qy 1724 aacttcctgcaaggttccgacgagatttcacggttacattcatcggagcagatgatttgg 1783  
Db 1745 agcttccatcaagaagttccttatgatttcatggaacattcataaacgccaatgatttgg 1804  
Qy 1784 cgaagcagg 1792  
Db 1805 caaatcagg 1813

RESULT 3  
AAD09399 standard; cDNA; 1839 BP.  
ID AAD09399  
XX AC AAD09399;  
XX DT 10-SEP-2001 (first entry)  
XX DE Vigna unguiculata neoxanthin cleavage enzyme, CPRD65 cDNA.  
XX KW Cowpea; neoxanthin cleavage enzyme; abscisic acid; ABA; herbicide;  
XX KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
XX KW plant growth protectant; Cowpea Responsive to Dehydration; CPRD65; ss.  
XX OS Vigna unguiculata.  
XX FH Location/Qualifiers  
XX CDS 1..1839  
FT /\*tag= a  
FT /product= "Vigna unguiculata CPRD65 protein"  
XX EP1116794-A2.  
XX PD 18-JUL-2001.  
XX PF 11-JAN-2001; 2001EP-0300218.  
XX PR 13-JAN-2000; 2000JP-0010056.  
XX PR 11-JAN-2001; 2001JP-0003476.

XX (RIKE ) RIKEN KK.  
PA Iuchi S, Kobayashi M, Shinozaki K;  
XX WPI; 2001-400081/43.  
XX P-PSDB; AAE04787.  
PT A DNA encoding a protein with a neoxanthin cleavage activity for  
PT producing transgenic plants with improved or decreased stress tolerance  
XX Claim 3; Page 53-56; 101pp; English.  
XX The invention relates to neoxanthin cleavage enzymes and their  
XX corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key  
XX role in endogenous abscisic acid (ABA) biosynthesis under drought stress.  
XX Neoxanthin cleavage enzyme is used for improving stress tolerance in a  
XX plant when expressed in a plant cell. The invention also relates to  
XX methods for increasing or decreasing stress tolerance in a plant by  
XX introducing the DNA into the plant, and a transgenic plant into which a  
XX neoxanthin cleavage enzyme is introduced. The improvement of stress  
XX tolerance in plants is useful, for example in plant breeding. Neoxanthin  
XX cleavage enzyme genes are useful for producing transgenic plants. An arid  
XX land can be improved by growing transformant weed for several years and  
XX then removing the weed by specifically lowering stress tolerance in the  
XX weed by inducing an inducible promoter. The present cDNA sequence encodes  
XX Vigna unguiculata neoxanthin cleavage enzyme, CPRD65 (Cowpea Responsive  
XX to Dehydration) protein. CPRD65 gene is isolated from cowpea plant.  
XX Sequence 1839 BP; 447 A; 595 C; 442 G; 355 T; 0 other;

Query Match 44.4%; Score 798.4; DB 22; Length 1839;  
Best Local Similarity 70.4%; Pred. No. 6.7e-248;  
Matches 1098; Conservative 0; Mismatches 456; Indels 6; Gaps 2;  
Qy 233 ccaagaatccaacataacagatgaattgttccagagagcgcgcgagcgttgg 292  
Db 281 ccaaccaccattacctcaaaatggaaactttctccgaaagccgtgcacggccttgg 340  
Qy 293 acgcgcgagggtttcttctgcagccacagagaagctacacccgtcttctaaaacgctg 352  
Db 341 acctggtcgaacggcgtgctctgcagagcgcaaacacccgtcccaaaacggcg 400  
Qy 353 atcctagtgtcagatcgccggaattttgtccggtgaatgaacagccctccgggcta 412  
Db 401 acccaggggtccaaatcgccgggaacttcgcccgtgcccggagcatgccgatacaag 460  
Qy 413 atctccggtggtcggaaaaacttcccgtatccatacaaggagtgatgtgcgaacggag 472  
Db 461 gactcccggttggcggaaaaatccccaaatgcatgacggtgacgtgcgcaacggcg 520  
Qy 473 ctaccacttcacgagccggtgacaggtcacacttcttcgacggagagcgttgatgttc 532  
Db 521 ccaatcgctctacagcctgtgcccgggaccacttcttcgacggcgagcgtatggtcc 580  
Qy 533 acgctgtcaaatcgaaacaggttcagctagctacgttgcggttttactcagactaac 592  
Db 581 acgctgtgaagtccagaaagcgcccgccagctacgctgcgcttcaccgagacgagc 640  
Qy 593 ggtttgttcaggaacgtcaattgggtgcacggttttccccaaagccatcgttgagttc 652  
Db 641 gtctctcgcgagagaataatctctagcccgccggtgttccccgaagggccatcgggagctcc 700  
Qy 653 acggccacacccggtattgcccgactcatgtattctacgcagagctgcagccggtatag 712  
Db 701 acggccactcggcatcgcggtctctcttctacgcgcggtctctctcgggtctcg 760  
Qy 713 tcgaccccgacacgggaacccggttagctaacgcggttgggttatattcaatggcggt 772  
Db 761 ttgatgggtcccgaggtggcggtggcgaacgcggtctcgtctacttcaacaaccacc 820



```
Qy 773 tattgctatgtcggagatgatttaccattaccagttcagatcactccaatgagatt 832
Db 821 tcttgccatgccgaagacgatttaccataacacgtgagaatcacccctaacgcgact 880
Qy 833 taataacccttggtcgttcgatttgcagacaaatagaatccacaatgattgccacc 892
Db 881 taaccaaccttggtcgttcgatttgcagacaaatagaatccacaatgattgccacc 940
Qy 893 cgaagtcgaccgggaatccggtgaacttctcgttttaagctacgagctcgtttcaaac 952
Db 941 cgaactggaccctgcgagcgactccacgctcagctacgactcattcagaagc 1000
Qy 953 cttaactaaatccttcgatttccacggaggaactaaatcacggagcgtcgagattc 1012
Db 1001 cttaactcaagtacttccgttctcccccgcgagcgtcgaactcccggaatcc 1060
Qy 1013 agcttgatcagcaacgagatgcagcatttcgcgattacagagaactcgtcgtcgaac 1072
Db 1061 cctgaaggagccaccatgatgcagcatttcgcataacggagaaatttcgctcgtcc 1120
Qy 1073 ctgaccagcaagctgtttcaagctgccggagatgatccgcggtggtctccggtggttt 1132
Db 1121 cgcaccagcaggtggtctcaactaacggagatgatcacccggcggtcccccgggtct 1180
Qy 1133 acgacaagacaaggtgcgaagattcgggatttttagacaataacgccgaagattcatoga 1192
Db 1181 acgacaagacaataactcaggttggatttctgcacaagaatgcgaaggcgcgaatg 1240
Qy 1193 acattaaagtggtatgatctcagatctctcgtctcctcattcctcgaacgcttgggaag 1252
Db 1241 cgtcgtggtggtacgcgcgcggtggttctcgtctcctcattcctcgaacgcttgggaag 1300
Qy 1253 agccgaaacagatgaagctcgtgagtggtcgtgattgatccgcggtggtctccggtggttt 1312
Db 1301 agccgaaacagagaggttgggtgatgggtcgtgattgatccgcggtggtctccggtggttt 1360
Qy 1313 tcaagagcttgacagaatctcaagagtgctcgtctgaaatccgctgaatcgaataa 1372
Db 1361 tcaacgaatgcgagagagatttgaagagcgtgctgcagagataaggtgaaactgagga 1420
Qy 1373 cgggtgaatcaactcgcctgcgacatctccaaagaaatcaacaaatcaacacctcgaag 1432
Db 1421 cgggaagctccactcggcgcgcctatctcgcgcgcg---acaagtgaaacctcgggaag 1477
Qy 1433 cagggtggtcacaagaaatcgtcggcgttaaaacaaatcgttacttgggttttag 1492
Db 1478 cggcatggtgaacagaacagctcggagaaagacccagtcggtatcggctcgg 1537
Qy 1493 cgaagcgttggtcctaaagctcagattcgtcgaagtgatctcactactcagagaagta 1552
Db 1538 cgaagccttgcccaagctcgggttgcgaagtgatttgcgtgagtgagggaagtga 1597
Qy 1553 aaaaacatcttcacggcataacgttaacggagagagcctcgttctcccgagagaag 1612
Db 1598 agaatacatgctatggagaagaagtgctcggggagcctcgttctctccccc---aacg 1654
Qy 1613 gagggagagaacgaagatacatcctcgttctcgttcacgacgagaagacatggaat 1672
Db 1655 gccaaaagaacagatgggtatattcggcattcgtgcacgacgagaagaatggaat 1714
Qy 1673 cgaagttacagatagtaacgcgttagcttagaggttgaagaacaggttaaacctcgt 1732
Db 1715 cgaagctgagattgtgaatgcccaaaatttaagctcgaagcttccatcaaacctccct 1774
Qy 1733 caaggttccgtacgatttcggtatcattcattcggagcagatatttggcgaagcgg 1792
Db 1775 ctctgttccctcacgggtttctatggaacttctcattcattcccaaggatttgggaacaag 1834
```

RESULT 4

AAD09394

ID AAD09394 standard; cDNA; 1752 BP.

XX

AC AAD09394;

XX 10-SEP-2001 (first entry)

XX Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED1 cDNA.

XX Neoxanthin cleavage enzyme; AtNCED1; abscisic acid; ABA; herbicide;  
stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
plant growth protectant; ss.

XX Arabidopsis thaliana.

FH Key Location/Qualifiers  
CDS 1..1752FT /\*tag= a  
FT /product= "Arabidopsis thaliana AtNCED1 protein"

XX EP1116794-A2.

XX 18-JUL-2001.

XX 11-JAN-2001; 2001EP-0300218.

XX 13-JAN-2000; 2000JP-0010056.

PR 11-JAN-2001; 2001JP-0003476.

XX (RIKE ) RIKEN KK.

XX Iuchi S, Kobayashi M, Shinozaki K;

XX WPI; 2001-400081/43.

DR P-PSDB; AAE04782.

XX A DNA encoding a protein with a neoxanthin cleavage activity for  
producing transgenic plants with improved or decreased stress tolerance

XX Claim 3; Page 18-22; 101pp; English.

XX The invention relates to neoxanthin cleavage enzymes and their  
corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key  
role in endogenous abscisic acid (ABA) biosynthesis under drought stress.  
Neoxanthin cleavage enzyme is used for improving stress tolerance in a  
plant when expressed in a plant cell. The invention also relates to  
methods for increasing or decreasing stress tolerance in a plant by  
introducing the DNA into the plant, and a transgenic plant into which a  
neoxanthin cleavage enzyme is introduced. The improvement of stress  
tolerance in plants is useful, for example in plant breeding. Neoxanthin  
cleavage enzyme genes are useful for producing transgenic plants. An arid  
land can be improved by growing transgenic plants for several years and  
then removing the weed by specifically lowering stress tolerance in the  
weed by inducing an inducible promoter. The present cDNA sequence encodes  
Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED1 protein.  
The AtNCED1 cDNA is obtained from an Arabidopsis plant-derived cDNA  
library using a cDNA of the cPRD65 (Cowpea Responsive to Dehydration)  
gene isolated from cowpea plant as a probe.

XX Sequence 1752 BP; 484 A; 396 C; 428 G; 444 T; 0 other;

Query Match 41.0%; Score 738.2; DB 22; Length 1752;  
Best Local Similarity 68.2%; Pred. No. 2.2e-228;  
Matches 1063; Conservative 0; Mismatches 478; Indels 18; Gaps 2;

Qy 244 aacactaaacagatgaattttccagagagcggcgccgagcgttggacgcgcgagag 303

Db 193 aatcctctccggcctaaacatcttcagaaagcggcggttgcgacgcgcgtgag 252

Qy 304 ggtttccttgcagcagagaagctacaccgcttctctaaacggctgacccagttgt 363

Db 253 cgtgcattaatctcacagcagcagagattctcacttccccaaacgcgtgacccagttgt 312

Qy 364 cagatcgcgcggaattttgtcccggtgaatgaacagcccgccggtccggtatcttcctcggtg 423

Db 313 cagatgcgggaattattcccggtaccggaatcttcggtccggcgaacctcaacgtc 372  
Qy 424 gtcggaataactcccgattccatcaaaagagtgatgctgcgaacggagctaaacctt 483  
Db 373 gaaggaaacatccctgactgactgacggtgttatatccgtaacggcgcaatccgatg 432  
Qy 484 cacgacgggtgacaggttcacactcttcacgagagcgttatggttcacgacgtcaaa 543  
Db 433 tttagcacaacagctggcgacatttaccgagagcgaatgggttcacgacgttaaa 492  
Qy 544 ttgaacacgggttcagctagctacgcttcgcggtttactcagactaacgggtttggttcag 603  
Db 493 ataaccaacgggttcagctagctacgacggtttacaaacacggagagattggttcag 552  
Qy 604 gaacgtcaattgggttcgacgggttttcccaaaagccatcggtgagcttcacggccacac 663  
Db 553 gaataacgattgggttcgacgggttttccgaagaacatcgcgagcttcacggctcaactcg 612  
Qy 664 ggtattgccgactcatctattacgcaagagctgcagcgggtatagtcgacccggca 723  
Db 613 ggaatcgacgttgatgctgtttacgcacgtggggtttgtggtctgatcaacaacaa 672  
Qy 724 caccgaacgggtgtagctaaacggcgggtttggtctatttcaatggcgggttattggctatg 783  
Db 673 aacggcgtcggagtagcaaacgcgggtttggtttactttaataaccggcttttagctatg 732  
Qy 784 tcggaggatgattacccttaccaggttcagatcactcccaatggagatttaaaacggtt 843  
Db 733 tcagaagacgatttaccgtaccaatataaaattactcaaacggcgatctccaaacggt 792  
Qy 844 ggtcgggttcgatttgatggacaattagaatcccaatgattgccaccggaagtcgac 903  
Db 793 ggaactacgatttcgacgggtcagtttaaaatccgcaatgtagctcacccgaaactggac 852  
Qy 904 ccggaatccggtggaactcttcggttaagctacgacgctcgtttcaaaagccttaccctaaa 963  
Db 853 ccggttacgaaggagcttcacgctgaagctacgacgctcgttaagaacacttaactgaaa 912  
Qy 964 tacttcgattctcaccggacggaactaaatcacgggacgtcgagattcagcttgatcag 1023  
Db 913 tacttcagattctccgacagcggcgttaaatccgcgggaattggagatcccgctcgaaact 972  
Qy 1024 ccaacgattgacagatttcggtattcagagaacttcgctcgtacgttcgacagcaa 1083  
Db 973 ccgacgattgattcagatttcgctataacggaattttggtgattccctgatcaacaa 1032  
Qy 1084 gtcgttttcaagctccgagatgatccggtgggttcggttcggttttacgacaagaac 1143  
Db 1033 gtcggttcaagctccgagatgattcccgtaaatcccggttggttttcgacggagaa 1092  
Qy 1144 aaggtcgaagattcgggatttttagacaaatcacccgaagatttcacgaacattaaagtgg 1203  
Db 1093 aaggtttccgattgggataatgcccgaagcgcgacgaagcttctcagataatctgg 1152  
Qy 1204 attgatgcagatgcttctgctccatctctggaacggttgggaagagccagaaacaa 1263  
Db 1153 gtgaactccggagacattctgtttctctctggaatgcatgggaatcccgagacg 1212  
Qy 1264 gatgaagtcgctgtagaggtccctgtagctccaccagactcaattttcaacgagctt 1323  
Db 1213 gaggagattggtgtagtcggatcgtgtagtgcgcggcggttcaatcttcaacgagaga 1272  
Qy 1324 gacgagaatctcaagagtgtcctgtctgaaatccgctgaaatctcaaaacccggtgaatca 1383  
Db 1273 gacgagcttgagaagcgttttgcggagatcaggataaaacctcagaacacgtaaaaacc 1332  
Qy 1384 actcgcctcagatcatctcaacgaagatcaaacagttcaacctcgaagcggatgctc 1443  
Db 1333 acgctcgtctcgttgggttaacgagat-----gtataatttagagattggtatggt 1386  
Qy 1444 aacagaacatgctcggcgcgttaaaaccaaattcgcttacttggctttgagcggcggtgg 1503

Db 1387 aaccggaacgggttaggaagaaaaaacccggttcgctgttttggctattgcttatccttg 1446  
Qy 1504 cctaaagtctcaggattcgtctaaagttagctactactgaggaagttaagaaaaatctt 1563  
Db 1447 ccaaaagtctccggttcgctaaaggtcgatctttgcaccggtgagatgaaaaatatt 1506  
Qy 1564 tacggcgataaacggttacgagagagcctctgtttctccccgg-----agaa 1611  
Db 1507 tacggcgggtgagaaatagcggcgaacccgtttcttccccggaactccggttaacggc 1566  
Qy 1612 ggagagaggaagcgaagatacatctctgtttctcacgacgagaagacatggaaa 1671  
Db 1567 gaagaaatgaagatgacggttatatatattgtcacgttcacgacgaagaacaaagaca 1626  
Qy 1672 tcggagttacagatagttaacgcggttagcttagaggttgaagcaacgggttaaaactccg 1731  
Db 1627 tcagaggtcagattataacgctgttaatttaaagttgaagctacgattaaactaccg 1686  
Qy 1732 tcaaggggtccgtacgatttcacggtacatttcacgagcggatgatttggcgaaagca 1790  
Db 1687 tctagatccggtatgggtttcattcggcacatttgggttcgaatgaactcgttgatca 1745

## RESULT 5

AA09400  
ID AAD09400 standard; cDNA; 1815 BP.

XX AAD09400;

DT 10-SEP-2001 (first entry)

XX Zea mays neoxanthin cleavage enzyme, VP14 cDNA.

XX Maize; neoxanthin cleavage enzyme; VP14; abscisic acid; ABA; herbicide;  
KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
KW plant growth protectant; ss.

XX Zea mays.

PH Key Location/Qualifiers  
CDS 1..1815  
FT /\*tag= a  
FT /product= "Zea mays VP14 protein"

XX EF1116794-A2.

XX 18-JUL-2001.

XX 11-JAN-2001; 2001EP-0300218.

XX 13-JAN-2000; 2000JP-0010056.

PR 11-JAN-2001; 2001JP-0003476.

XX (RIKE ) RIKEN KK.

XX Iuchi S, Kobayashi M, Shinozaki K;

XX WPI; 2001-400081/43.

DR P-PSDB; AAE04788.

XX A DNA encoding a protein with a neoxanthin cleavage activity for  
PT producing transgenic plants with improved or decreased stress tolerance

XX Claim 3; Page 60-64; 101pp; English.

XX The invention relates to neoxanthin cleavage enzymes and their  
CC corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key  
CC role in endogenous abscisic acid (ABA) biosynthesis under drought stress.  
CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a  
CC plant when expressed in a plant cell. The invention also relates to  
CC methods for increasing or decreasing stress tolerance in a plant by  
CC introducing the DNA into the plant, and a transgenic plant into which a





QY 1794 cgtgtga 1800  
|  
Db 1728 ttcttga 1734

# RESULT 7

AAC56678  
ID AAC56678 standard; DNA; 443 BP.

XX  
AC AAC56678;

XX 25-JAN-2001 (first entry)

XX Eucalyptus grandis transcription factor DNA sequence #549.

XX Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
KW homeodomain zipper; LIM domain; AP2; ERBs; zinc finger domain;  
XX type 2 Cys2His2; CCAAT box element; MYB; ss.

XX Eucalyptus grandis.

OS WO200053724-A2.

PN 14-SEP-2000.

PD 09-MAR-2000; 2000WO-US06112.

PF 11-MAR-1999; 99US-026513.

PR 18-AUG-1999; 99US-0149485.

XX (GENE-) GENESIS RES & DEV CORP LTD.

PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX Wood M, McGrath A, Shenk MA, Glenn M;

XX WPI; 2000-579369/54.

DR New isolated polynucleotide encoding a plant transcription factor for

XX producing a plant e.g. a woody plant, preferably eucalyptus or pine,

XX having modified gene expression or modified activity of a polypeptide

XX Claim 1; Page 490; 747pp; English.

XX The present invention relates to novel plant transcription factors from  
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
CC sequence for one such transcription factor. The transcription factor may  
CC be used to produce a plant having modified gene expression such as a  
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
CC mahogany species or to modify the activity of a polypeptide in a plant.  
CC The transcription factors of the present invention are members from the  
CC following families of regulatory proteins: bZIP, bZIP family of G-box  
CC binding factors, basic helix-loop-helix zipper,  
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
CC and ERBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
CC and MYB.

XX Sequence 443 BP; 71 A; 177 C; 117 G; 77 T; 1 other;

Query Match 12.8%; Score 229.8; DB 21; Length 443;  
Best Local Similarity 69.9%; Pred No. 8.5e-64;  
Matches 309; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 620 gaccggtttcccaagccatcggtgagcttcacggccacacccggtatgccgactca 679

Db 1 gcccgcttcncaagccatcggtgagcttcacggccacacccggtatgccgactca 60

QY 680 tgcatttaccgagctgacgcggtatagtgcaccccgacacggaaacgggtgag 739

|||||

Db 61 tgcatttaccgagctgacgcggtatagtgcaccccggaatggcatgggctg 120  
QY 740 ctacgcgggtttgctctatttcaatggccggttattggctatgctgaggatgatttac 799  
|  
Db 121 cgaacgcggctgctgacttcgacggccacctctcgcatgtccgaggacacctcc 180  
QY 800 cttaccagttcagatcactcctccaatggagattttaaaaacggttggttcgattttg 859  
|  
Db 181 cctaccagtgcgctcagcgctccggcagctcgagacgtcgccgctacgacttcg 240  
QY 860 atggacaattagattcccaatgattcccccacccgaaagtcgacccggaatccggtgaac 919  
|  
Db 241 ccggccagctgacttcgcatgacgccccacccggaagtcgacccggttcggcgaga 300  
QY 920 tcttcgcttaagctacgacgtcgtttcaagccttacctaaaatacttccgattctcac 979  
|  
Db 301 tgttcgctcagctacgacgtcgtccggaagcgtacctcaagtacttcgattctcca 360  
QY 980 cgaacggaactaaatcacggacgtcgagattcagcttgatcagcgaacgattgacgacg 1039  
|  
Db 361 aggaacgcgagaagtcctcccgacgtcgagatccccctggtgagccgacctgatgcacg 420  
QY 1040 atttcgcatcagagaaatt 1061  
|  
Db 421 atttcgcatcagagaaatt 442

# RESULT 8

AAC82706/C

ID AAC82706 standard; DNA; 492 BP.

XX AAC82706;

XX 15-MAR-2001 (first entry)

XX Rice abscisic acid synthesis associated DNA SEQ ID NO 5.

XX Rice; abscisic acid synthesis; VP14-like gene; drought resistance; ds.

OS Oryza sativa.

PN WO200071727-A1.

PD 30-NOV-2000.

PF 25-MAY-1999; 99WO-JP02734.

XX 25-MAY-1999; 99WO-JP02734.

XX (NORQ) JAPAN MIN AGRIC FORESTRY & FISHERIES.

PA (BIOO-) BIO-ORIENTED TECHNOLOGY RES ADVANCEMENT.

PI Hirochika H, Sakamoto K;

XX WPI; 2001-032042/04.

XX Oligonucleotide encoding gene for regulating abscisic acid synthesis in  
PT plants, useful for constructing e.g. genetically-modified rice with  
PT drought resistance and ear-germination resistance

XX Disclosure; Page 50; 55pp; Japanese.

XX This invention describes a novel polynucleotide sequence (I) which  
CC encodes a protein capable of regulating the synthesis of abscisic acid.  
CC The invention also describes (1) an oligonucleotide encoding a protein  
CC of a gene obtained by controlling the expression of a VP14-like gene;  
CC (2) a vector containing the oligonucleotide ligated operably to the  
CC regulation sequence; (3) a plant transformed with the vector; and  
CC (4) a method for regulating abscisic acid synthesis in a plant including  
CC the transfer of the above oligonucleotide to it. The gene is useful for  
CC constructing drought resistant rice.

XX Sequence 492 BP; 90 A; 126 C; 186 G; 88 T; 2 other;

```
Query Match      11.3%; Score 204; DB 22; Length 492;
Best Local Similarity 69.7%; Pred. No. 2.1e-55;
Matches 304; Conservative 0; Mismatches 130; Indels 2; Gaps 2;

QY 627 tttcccaaaagccatcggtgagcttcacggccacacc-ggtattgccgactcatgttat 685
DB 435 TTTTCTAAGCGTAGGTAGCTCCATGGCCACTCCGGGCATCGCGCCTTGCTCTGT 376
QY 686 tctacgccagagctgacgggtatagtcgaccgggcacacgggaacccggtgtagtaacg 745
DB 375 TCTAGCGCGCGCGC-CTCGGCTCTCTCGACCCGTCAACGGCACCGGCGTGCGCCAAG 317
QY 746 ccggtttgggtatttcaatggccggttatggcttatctcgagagatgatttaccttacc 805
DB 316 CCGGCTCATCTACTTCAACGGCAGGCTCTCGCCATGTCGGAGGACGACCTCCCTTACC 257
QY 806 aagttcagatcactcccaatgagatttataaaacggttggctggttcgattttgatggac 865
DB 256 AGTTCGGCTTACCGCCGACGGCGACCTCGAGACCTCGGCGCTACGACTTCGACGGGC 197
QY 866 aattagaatccacaatgattgccaccgcgaagtgcacccgggaatccggtgaactcttcg 925
DB 196 AGCTCGGTGCGCCATGATCGGCACACCCCAAGCTCGACCCGCGCACCGAGAGTCCACG 137
QY 926 ctttaagtacagcgtggttcaagccttaccataaaatacttcgatttctcacggagc 985
DB 136 CGCTCAGCTACGACGTGATCAAGAAGCGGTACTCTCAAGTACTTCTTACTTCGCGCCGACG 77
QY 986 gaactaaatcacggacgtcagatttcagcttgatcagcaacgatgatgcagatttcg 1045
DB 76 GCACCAATCGGCGACGTGAGATCCCGCTCGACCCACCCACCATGATGATCCAGACTTCC 17
QY 1046 cgattacagagaactt 1061
DB 16 GAATTACTGAGTCTT 1

RESULT 9
AAC56695
ID AAC56695 standard; DNA; 372 BP.
XX
AC AAC56695;
XX
DT 25-JAN-2001 (first entry)
XX
DE Eucalyptus grandis transcription factor DNA sequence #566.
XX
KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
KW type 2 Cys2His2; CCAAT box element; MYB; ss.
XX
OS Eucalyptus grandis.
XX
FN WO200053724-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-US06112.
XX
PR 11-MAR-1999; 99US-0266513.
PR 18-AUG-1999; 99US-0149485.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Wood M, McGrath A, Shenk MA, Glenn M;
XX WPI; 2000-579369/54.
XX
```

New isolated polynucleotide encoding a plant transcription factor for producing a plant e.g. a woody plant, preferably eucalyptus or pine, having modified gene expression or modified activity of a polypeptide

Claim 1; Page 494; 747pp; English.

The present invention relates to novel plant transcription factors from *Eucalyptus grandis* or *Pinus radiata*. The present sequence is the coding sequence for one such transcription factor. The transcription factor may be used to produce a plant having modified gene expression such as a woody plant e.g. a *eucalyptus*, pine, acacia, poplar, sweetgum, teak, or mahogany species or to modify the activity of a polypeptide in a plant. The transcription factors of the present invention are members from the following families of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements and MYB.

Sequence 372 BP; 60 A; 148 C; 101 G; 63 T; 0 other;

Query Match 10.4%; Score 188; DB 21; Length 372;  
Best Local Similarity 69.1%; Pred. No. 2.8e-50;  
Matches 257; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 639 catcggtgagcttcacggccacacccggtattgccgactcatgtattctacgccagac 698  
DB 1 catcgcgagctccacggccactccgcatcgcggtcatgctcttctacgcccgac 60  
QY 699 tgcagcgggtatagtcgacccggcacacgggaacgggtgtagctaacccgggttggctta 758  
DB 61 cctcttcggcctcgctgacaccgggaatggcatggcggtcggaacccggcctcgtga 120  
QY 759 ttccaatggcggttattggtatgctcgaggatgattaccttaccaggtcagatcac 818  
DB 121 cttegagggccactctcgcgatgctcgaggagacactccctaccacgtgcgcgtcac 180  
QY 819 tcccaatggagatttataaaacccgttgcgtcggttcgatttgcggaacattagatccac 878  
DB 181 gcgtccggcgacctcgagaccgtcgccgctacgacttcgcccggcagctcgactctcc 240  
QY 879 aatgattgccccccgaaagtcgacccgggaatccgggtgaactcttcgctttaaagtcga 938  
DB 241 gatgatcgccacccgaaagatcgaccgggttcggcgagatgctccctcagctacga 300  
QY 939 cgtcgtttcaaaagccttaccataaaacttccgatttccacggacggaaactaaatcac 998  
DB 301 cgtcgtccggaagcgtacctcaagtacttccgatttcccaaggacggcgagaagtcccc 360  
QY 999 ggacgtcgagat 1010  
DB 361 cgacgtcgagat 372

RESULT 10  
AAC56548  
ID AAC56548 standard; DNA; 325 BP.  
XX  
AC AAC56548;  
XX  
DT 25-JAN-2001 (first entry)  
XX  
DE Eucalyptus grandis transcription factor DNA sequence #419.  
XX  
KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;  
KW type 2 Cys2His2; CCAAT box element; MYB; ss.  
XX  
OS Eucalyptus grandis.

XX WO200053724-A2.  
 XX PD 14-SEP-2000.  
 XX PF 09-MAR-2000; 2000WO-US06112.  
 XX PR 11-MAR-1999; 99US-0266513.  
 XX PR 18-AUG-1999; 99US-0149485.  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
 XX PI Wood M, McGrath A, Shenk MA, Glenn M;  
 XX WPI; 2000-579369/54.  
 XX New isolated polynucleotide encoding a plant transcription factor for  
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
 PT having modified gene expression or modified activity of a polypeptide  
 PT  
 XX  
 PS Claim 1; Page 461; 747pp; English.  
 XX The present invention relates to novel plant transcription factors from  
 CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
 CC sequence for one such transcription factor. The transcription factor may  
 CC be used to produce a plant having modified gene expression such as a  
 CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
 CC mahogany species or to modify the activity of a polypeptide in a plant.  
 CC The transcription factors of the present invention are members from the  
 CC following families of regulatory proteins: bZIP, bZIP family of G-box  
 CC binding factors, basic helix-loop-helix zipper,  
 CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
 CC and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
 CC and MYB.  
 XX  
 SQ Sequence 325 BP; 45 A; 128 C; 100 G; 52 T; 0 other;

Query Match 9.5%; Score 171.8; DB 21; Length 325;  
 Best Local Similarity 71.2%; Pred. No. 4.6e-45;  
 Matches 227; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 451 ggagtgtatgtgcgaacggagctaacacccttcacgagccggtgacaggtcaccacttc 510  
 Db 3 ggggtgtacgtccgaacggcgccaccgcctccacgagccggtgcggggcaccacttg 62  
 Qy 511 ttcgacgagacggatgggttcacgcccgtcaaatcgaacacgggttcagctacgct 570  
 Db 63 ttcgacggcgaacgcatgatccacgcgtccggttctccggggtcagtgagctacgcc 122  
 Qy 571 tgcgggtttactcagactaacgggtttgttcaggaaactcaattgggtcagccggttttc 630  
 Db 123 tgcgggttcacgagacgaacacgctgatccaggaaacggggtccgcccgcctcttc 182  
 Qy 631 cccaaagcctcgttgagcttcacggccacacacggatattgcccgactcgtattctctac 690  
 Db 183 cccaaggccatcggcagctccacggccactccggcgtcgcgggtcgtctctctac 242  
 Qy 691 gccagagctgcagccggtatagtcacccgggcacacggaacgggtgtagctaacgccggt 750  
 Db 243 gcccgccgctcttcggctcgtcgcacacccgaatggcgtggcggtcgcgaacgcgggc 302  
 Qy 751 ttggtctatttcaatggcc 769  
 Db 303 ctctgtacttcgacggcc 321

RESULT 11  
 AAC42989  
 ID AAC42989 standard; DNA; 1788 BP.  
 XX

AC AAC42989;  
 XX 17-OCT-2000 (first entry)  
 DT XX Arabidopsis thaliana DNA fragment SEQ ID NO: 37611.  
 DE XX Hybridisation assay; genetic mapping; gene expression control;  
 XX protein identification; signal transduction pathway;  
 KW metabolic pathway; promoter; termination sequence; ss.  
 KW Arabidopsis thaliana.  
 OS Arabidopsis thaliana.  
 XX EP1033405-A2.  
 PN PD 06-SEP-2000.  
 XX 25-FEB-2000; 2000EP-0301439.  
 PF 25-FEB-1999; 99US-0121825.  
 XX 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 18-MAY-1999; 99US-0134370.  
 PR 19-MAY-1999; 99US-0134768.  
 PR 20-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.

PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	24-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142300.	PR	22-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145152.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.
PR	03-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161992.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0162142.
PR	04-AUG-1999;	99US-0147302.	Query Match 9.4%; Score 169.6; DB 21; Length 1788; Best Local Similarity 50.3%; Pred. No. 6.5e-44; Matches 709; Conservative 0; Mismatches 634; Indels 69; Gaps 9;		
PR	05-AUG-1999;	99US-0147192.			
PR	05-AUG-1999;	99US-0147260.	Qy	426	cggaaaaacttcccgattccatcaaggagtgatgtgcgcaacggagctaacccacttca 485
PR	06-AUG-1999;	99US-0147303.	Db	393	cggcaactcttccactgtcacttaacgcgctttacatccgtaacggtccaaatccacagtt 452
PR	06-AUG-1999;	99US-0147416.	Qy	486	cgagcgggtgacaggtcaccaacttcttcgacgagacggtatggttcacgcgtaatt 545
PR	09-AUG-1999;	99US-0147493.	Db	453	tctccctcggtggtccttaccatctcttcgacggtgacggtatggttcacgccccataaaat 512
PR	10-AUG-1999;	99US-0147935.	Qy	546	cgaacacggttcagctagctacggttgcgggtttactcagactaacccggtttgttcagga 605
PR	11-AUG-1999;	99US-0148171.	Db	513	ccacaacggtaaagccactctctgtagcagatacgtcaagacttataataacaacgctga 572
PR	11-AUG-1999;	99US-0148319.	Qy	606	acgtcaattgggtcgaccggttttccccaagccatcggtagcttcacggccacacgg- 664
PR	12-AUG-1999;	99US-0148341.	Db	573	gaaacaaacggagctccggttatgcctaacgtgttttccggattcaacggtgtacgggc 632
PR	13-AUG-1999;	99US-0148565.			
PR	16-AUG-1999;	99US-0148684.			
PR	17-AUG-1999;	99US-0149368.			
PR	18-AUG-1999;	99US-0149175.			
PR	20-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149722.			
PR	20-AUG-1999;	99US-0149723.			
PR	23-AUG-1999;	99US-0149929.			
PR	23-AUG-1999;	99US-0149929.			
PR	23-AUG-1999;	99US-0149930.			
PR	25-AUG-1999;	99US-0150566.			



```
QY 665 --gtattgccgaactcatgctattctacgccagagctgacgcggtatagtcgacccgcg 722
Db 633 gtcagtagctcgggacttttaacgcgacgtagggttttaacgcgacagataatccgat 692
QY 723 acacggaacgggtgtagtaaacgcggtttggtctatttcaatggcgggttattggctat 782
Db 693 taacggcattggttagtaataacaagtcagcttcttcagtaacgcgtctcttggcttt 752
QY 783 gtccggggatgattttaccttaccaaagtccagatcactcccaatggagatttataaacct 842
Db 753 aggtgaatctgattttaccttaccgcgcgtattaaacgaatcaggagatattgaaacgat 812
QY 843 tgggtcgggttcgatttgtaggacaattagaatcccaatgattgcccacccgaaagtcga 902
Db 813 cggacggtacgatttcgacgggaattagcagatgagtagacagctcatcctaaaacga 872
QY 903 ccggaatccggtgaactcttcgcttttaagctacgacgctggttcaaaagcccttacctaa 962
Db 873 tccaataacggagaaactttcgctttccggtacggtccggttcca---ccggttttaac 929
QY 963 atacttcgattctccacggaggaactaaatcacgcgacgctcgaga---ttcagcttga 1019
Db 930 atatttcggtttgattccgcgggaaaaaacaagagacggtccgatatctcgatgac 989
QY 1020 tcagccaacgatgatgcacgatttcgcgattacagaaacttcgctcgtaacctgacca 1079
Db 990 gtcctcgctgttctccatgacttcgcgacacgaaacgacgacgttttcgcgagat 1049
QY 1080 gcaagtcg-----ttttcaagctgcggagatgacccggtgggtctcccggtggt 1130
Db 1050 tcagcttggcatgaggatgaacatgtggattgggtcttcgaaggcgttctccggttgg 1109
QY 1131 ttacgacaagaacagggtcgcaagattcgggattttagacaatacgcggaagattcatc 1190
Db 1110 tactgataacggaaaaactccaagcttggagtgattcttaagtagcgcggagagatgc 1169
QY 1191 gaacattaagtgattgatgctccagatgcttctcgtctccatctctcgaaacgcttgga 1250
Db 1170 ggagatgaatggttcgaagtcctggattcattcaatcattcagctattaatgcttgg 1229
QY 1251 agagccagaacacagatgaagtcgctgctgataagggtcctgctgactccaccagactca 1310
Db 1230 tgaagatgatgaaacagcgtcgtttgttgcacccaataattatgtcgtgattgaacata 1289
QY 1311 tttaacgagctgcagagaatctcaagagtgctcgtctgataatccgctgaatctcaa 1370
Db 1290 tttagagagatggatctgttcc---atgcttgggtggagaagggaagatcgatcgt 1346
QY 1371 aaccggtgaatacaactccgcgtccgatactctccaaacgaagatcaacaagtcacaactcg 1430
Db 1347 caccgggattgtgagacgtcatccgactcagcgagg-----aatctcga 1391
QY 1431 agcaggatgggtcaacagaaacatcctcgccgttaaaaccaaattcgcttacttgcttt 1490
Db 1392 ttctcgtgataataatccggcgttctcgggagatgtagcaggtacgtttacgcggcgat 1451
QY 1491 agccagcgtggtcctaagtcaggatcgcgtctaaagtgtatctcactactcggagaagt 1550
Db 1452 tggagatccgatgcgaagatcctcggtggtggaagctgtatgtctaaaggagatcg 1511
QY 1551 taagaataat-----ctttacgcggataacccttaccgaggagagacccct 1595
Db 1512 ggaatgattgacggtggcccgatgagatgacggtttcaggttttaccggcggaacccgt 1571
QY 1596 gtttctcccgaga-----aggaagagagaagacgaagatacatctctcg 1643
Db 1572 ttctcgtagtagggatccttggttaatccggagcgaggagagatgaggtatgtgtggtgac 1631
QY 1644 ttctcgtcacgacgagaaacatggaatcgaggttcagatagtttaacgcgcttagctt 1703
Db 1632 gtagtgcacgatgaagtagctggagaatcgaagttcttggtgtaggacgctaaatcgcc 1691
QY 1704 agaggttgaa-----gcaacggtttaaacttccgtccgaaggttccgtagcggatttcacgg 1757
```

```
Db 1692 ggagcttgaatacgtcgccgcgctgaggttcgcggaagggttccgtagcggattccatg 1751
QY 1758 tacattcatcggagcgcgatgatttggcgaaagc 1789
Db 1752 gttattgtcaaggaaagtgcaccttaataagc 1783

RESULT 12
AAD09395
ID AAD09395 standard; cDNA; 1788 BP.
XX
AC AAD09395;
XX
DT 10-SEP-2001 (first entry)
XX
DE Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED2 cDNA.
XX
KW Neoxanthin cleavage enzyme; AtNCED2; abscisic acid; ABA; herbicide;
KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;
KW plant growth protectant; gs.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
CDS 1..1788
FT /*tag= a
FT /product= "Arabidopsis thaliana AtNCED2 protein"
XX
FN EP1116794-A2.
XX
PD 18-JUL-2001.
XX
PF 11-JAN-2001; 2001EP-0300218.
XX
PR 13-JAN-2000; 2000JP-0010056.
PR 11-JAN-2001; 2001JP-0003476.
XX
PA (RIKE ) RIKEN KK.
XX
PI Iuchi S, Kobayashi M, Shinozaki K;
XX
WPI; 2001-400081/43.
DR P-PSDB; AAE04783.
XX
PT A DNA encoding a protein with a neoxanthin cleavage activity for
producing transgenic plants with improved or decreased stress tolerance
XX
XX
Example 10; Page 25-29; 101pp; English.
XX
The invention relates to neoxanthin cleavage enzymes and their
corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key
role in endogenous abscisic acid (ABA) biosynthesis under drought stress.
Neoxanthin cleavage enzyme is used for improving stress tolerance in a
plant when expressed in a plant cell. The invention also relates to
methods for increasing or decreasing stress tolerance in a plant by
introducing the DNA into the plant, and a transgenic plant into which a
neoxanthin cleavage enzyme is introduced. The improvement of stress
tolerance in plants is useful, for example in plant breeding. Neoxanthin
cleavage enzyme genes are useful for producing transgenic plants. An arid
land can be improved by growing transformant weed for several years and
then removing the weed by specifically lowering stress tolerance in the
weed by inducing an inducible promoter. The present cDNA sequence encodes
Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED2 protein.
The AtNCED2 cDNA is obtained from an Arabidopsis plant-derived cDNA
library using a cDNA of the cPRD65 (CowPea Responsive to Dehydration)
gene isolated from cowpea plant as a probe.
XX
Sequence 1788 BP; 444 A; 444 C; 418 G; 482 T; 0 other;
```

Query Match

9.4%; Score 169.6; DB 22; Length 1788;

Best Local Similarity 50.2%; Pred. No. 6.5e-44;		Matches 709; Conservative 0; Mismatches 634; Indels 69; Gaps 9;	
Qy	426	cggaataattccgattccatcaagaagtgatgtgcgcaacggagctaacccacttca	485
Db	393	cggaactttccactgtcacttaacggcgcttaacccgtacccgtacccacacagt	452
Qy	486	cgagccggtgacaggtcaccactcttcgacggagacgggtatggttcacgcgcgaatt	545
Db	453	tctccctcgttggttcaccatctcttcgacggcgaggtatggttcacgcacataaaat	512
Qy	546	cgaaacaggttcagtagactacgtttgpcggtttactcagactaacgggtttgttcagga	605
Db	513	ccaaacaggttaaacccactctctgacagatacgtcaagacttataatacaacgctga	572
Qy	606	acgtcaattgggtcgacgggttttcccacaaagccatcggtgagcttcacggccacacg-	664
Db	573	gaacaaacggagctccggttatgcttaacgtgttttccggattcaacgggtgaacggc	632
Qy	665	--gtattcccgactcatgctattctacccagagctgcacccggtatagtcgacccggc	722
Db	633	gtcagtagctcgtggagctttaacggcagctagggttttaaccggacagtataatccggt	692
Qy	723	acacggaaacgggttagcttaacgcgggtttgtctatttcaatggccggttatgggtat	782
Db	693	taacggcaattggttagcttaacaaagctcagctagctttctcagtaacgcgtctcttgttt	752
Qy	783	gtcggagagatgatttaccttaccagttcagatcactcccaatggagatttaaaaacgt	842
Db	753	aggtgaatctgatttacctacgcgcgtccgatttaacggaatcaggagatattgaagat	812
Qy	843	tggctggttcgattttgatgacataagataacacatgattcccaaccgaaagtcga	902
Db	813	cggacggtacgatttcgacgggaatagcgtagctagctacagctcatctctaaaccca	872
Qy	903	cccggaaatccggtgaactcttcggttaagctacgaogctgtttcaagccttacctaaa	962
Db	873	tccaataacggagaaacttcggttccggtacggtccggttcca--cgtttttaac	929
Qy	963	atacttcgattctcacggcagcgaactaaatacccggaactcaga---ttcagcttga	1019
Db	930	atatttcggtttgattccgcgggaaacaaagacggttccgatatattctcgatgac	989
Qy	1020	tcacccaacgtagatgacgatttcgagattacagaaacttcgctcgtacactgacca	1079
Db	990	gtctccctcgtttctccatgacttcgagatcacgaaacgtcacgcgattttcgcagagat	1049
Qy	1080	gcaagtcg-----ttttcaagctccggagatgacccggttggttcctccggtggt	1130
Db	1050	tcagcttggeatgaggatgaacatgttgatttggttcgaaaggtggttcctccggttg	1109
Qy	1131	ttacgacaagaacaggttcgcaagattcgggatttttagacaaataccgccgaagattcatc	1190
Db	1110	tactgataacggaaacacccaagcttgagtgattctcaagtaacgcgcggagatgagtc	1169
Qy	1191	gaacattaagtgagattgatgctcagatggttctgctccatctcttggaacgcttggga	1250
Db	1170	ggagatgaattggttcgaagttcctggattcaatatcattcacgctattaatgcttggga	1229
Qy	1251	agaccagaacacagatgaagtcgctggtatagggttcctgtatgactccaccagactcaat	1310
Db	1230	tgaagatgatggaacacgctcgttttgatgcaccgaatattatgctgattgaacatac	1289
Qy	1311	tttcaacgagctgcagcagaatctcaagagtgctcgtctctgaaatccgcctgaatctcaa	1370
Db	1290	tttagagagatgagatcgtggttc---atgctttggtggagaaggtgaagatcgatctcgt	1346
Qy	1371	aacgggaaatcaactccgctcgcgactatctccaacgaagatcaacaaagtcgaacccga	1430
Db	1347	caccgggatttgagacgtcatcgatcgcgaggg-----aatctcga	1391
Qy	1431	agcagggatggtcaacagaacatgctcggccgtaaaaccaaattcgcttacttggcttt	1490
Db	1392	tttcgctgtgattaataatccggctttctcgggagatgtagcaggtagcttttacgcggcgat	1451
Qy	1491	agccgaacgctggcctaataatctcagattcgtataaagttgattctcactactggagaagt	1550
Db	1452	tggagatccgatccgaagatcccggttggtggaagcttgatgctctaaaggagatcg	1511
Qy	1551	taagaaacat-----ctttacggcgataaacggttacgggagagagcctct	1595
Db	1512	ggatgattgtacggtggcccgtagaatgtacggttcagggtgttacggcgggagaacggtt	1571
Qy	1596	gtttctcccgggaga-----aggagagaggaagacgaagatacatcctctg	1643
Db	1572	tttcgtagctaggatcctgtaatccggaggcgaggagatgattggttatgfggtgac	1631
Qy	1644	tttcgttcacgacgagaagacatggaaatcggagttcacagatagttaacgcggttagctt	1703
Db	1632	gtatgttcacgataagtgactgagaaatcgaagtttctggtgatggacgctaaatcgcc	1691
Qy	1704	agaggttgaa-----gcaacggttaaaccttcggtcaaggtttccgtacggatttcacgg	1757
Db	1692	ggagcttgaaatcgtccgcccggaggttgcgcggaagggttccgtacggattccatg	1751
Qy	1758	tacattcatcggagccgagatgatttggcgaagc	1789
Db	1752	gttattgtcaaggaaagtgccttaataagc	1783
RESULT 13			
AAAF77206			
ID	AAF77206 standard; cdNA; 1950 BP.		
XX	AC AAF77206;		
XX	AC AAF77206;		
DT	16-MAY-2001 (first entry)		
XX	cdNA encoding sunflower neoxanthin cleavage enzyme (NCE).		
DE	Defence-related signalling gene; sunflower; neoxanthin cleavage enzyme;		
KW	NCE; amino acid permease; AAP; glutamic acid rich protein; GRP;		
KW	pathogen resistance; abscisic acid metabolism; ss.		
XX	Helianthus annuus.		
XX	Location/Qualifiers		
PH	1..1632		
FT	/*tag= a		
FT	/product= "NCE"		
FT	/note= "Neoxanthin cleavage enzyme"		
FT	/partial		
XX	WO200112801-A2.		
PN	22-FEB-2001.		
XX	17-AUG-2000; 2000WO-US22961.		
XX	18-AUG-1999; 99US-0149656.		
PR	23-MAY-2000; 2000US-0206405.		
XX	(PION-) PIONEER HI-BRED INT INC.		
PA	(CURA-) CURAGEN CORP.		
PA	Bidney DL, Crasta OR, Hu X, Lu G;		
XX	WPI; 2001-211215/21.		
DR	P-PSDB; AAB72303.		
XX	Novel isolated defence-related signalling gene isolated from sunflower		
PT	encoding neoxanthin cleavage enzyme, amino acid permease or glutamic		
PT	acid-rich protein useful for increasing resistance of plant to a		
PT	pathogen		
XX	Claim 1; Page 94-97; 135pp; English.		
PS			

XX This invention relates to defence-related signalling genes isolated from  
 CC the sunflower (*Helianthus annuus*). The genes encode a neoxanthin cleavage  
 CC enzyme (NCE), an amino acid permease (AAP) and a glutamic acid rich  
 CC protein (GRP). The signalling gene is useful for increasing the  
 CC resistance of a plant to a pathogen such as fungus, virus, bacterium,  
 CC nematode or insect (e.g. European corn borer), preferably incorporating a  
 CC *Sclerotinia* spp., *Phoma* spp., or *Phomopsis* spp., by stably incorporating a  
 CC construct containing the gene into the genome of the plant. The gene is  
 CC useful for regulating gene expression in a plant, in response to a  
 CC stimulus such as infection with a pathogen, damage from a pathogen,  
 CC hydrogen peroxide, jasmonic acid, methyl jasmonate, salicylic acid,  
 CC oxalic acid or expression of a gene encoding oxalic acid oxidase. The  
 CC genes are also useful for stem-preferred regulation of gene expression in  
 CC a plant. The genes are useful in agriculture, particularly in the  
 CC breeding of crop plants with improved agronomic traits, for modifying  
 CC abscisic acid (ABA) metabolism and for modifying amino acid transport and  
 CC content in plants. The present sequence represents cDNA encoding the  
 CC sunflower neoxanthin cleavage enzyme (NCE).  
 XX  
 SQ Sequence 1950 BP; 535 A; 456 C; 453 G; 506 T; 0 other;

Query Match 9.4%; Score 169.2; DB 22; Length 1950;  
 Best Local Similarity 51.0%; Pred. No. 9.2e-44;  
 Matches 480; Conservative 0; Mismatches 453; Indels 9; Gaps 3;

QY 347 cggctgacccctagctcagatcgccggaatttgcctccggtgaatgaacagccgctcc 406  
 Db |||||  
 QY 308 cagttgatccaaaacagctttgtctgataactttccacgggagcaactccctccga 367  
 Db |||||  
 QY 407 ggcgtaatcttcgggtggcggaaaacttcccgattccatcaaaagagtgatgcgca 466  
 Db |||||  
 QY 368 ctgactgtgaagtcatcgaggccacactcccaagttgccttgacggtgttactccgta 427  
 Db |||||  
 QY 467 acggagctaaccaacttcacgagcgggtacaggtcaccaactttccagcagacgcta 526  
 Db |||||  
 QY 428 atggtccgacccgcaattccctccgcgaggaccctaccacctcttcgatgagcgta 487  
 Db |||||  
 QY 527 tgggtcacgcccgtcaaatccgaacaggttcagctagctacgcttgcgggtttactcaga 586  
 Db |||||  
 QY 488 tgcctcagctattcgatctccatggaagagcttcgtatgtagcgcgatacatcaaaa 547  
 Db |||||  
 QY 587 ctacccggtttgttcaggagacgtcaattgggtcgaccgggttttcccaagccatcggtg 646  
 Db |||||  
 QY 548 cacaataattcattagagaagaagcgggattcccaattattccaaacgtgtttcag 607  
 Db |||||  
 QY 647 agcttcacggccacacccggtattgcccagctcatgctattc---tacgcagagctgcag 703  
 Db |||||  
 QY 608 ggttaagtgtgactgctctcagctcgcagtcgcagtcagctcgcgagctcgcggttttg 667  
 Db |||||  
 QY 704 ccggtatagtcgacccggcacacggaacgggtgtagtaaacccggtttgtctatttca 763  
 Db |||||  
 QY 668 ctggacaatttgaccccaaaaaggtattggtctgccaataccagctcggcctttttg 727  
 Db |||||  
 QY 764 atggccggttattggtctatgctggaggatgatttacccttaccagttccagatcactcca 823  
 Db |||||  
 QY 728 gcaacagacttttgcctcggagagtcggtatcccatatccgtccaactagcgcgcg 787  
 Db |||||  
 QY 824 atggagatttaaaacccgttgctcgttgcgatttttgatggacaattagatcccaatga 883  
 Db |||||  
 QY 788 acggtgacatagtcacccgctcggacgtgaggaacttcgacggcaaatattcatgagcatga 847  
 Db |||||  
 QY 884 tggccacccgaaagtccagccggaatccggtgaactcttcgctttaagctacgagctcg 943  
 Db |||||  
 QY 848 ccgctcaccacaaaatccgatccagtaacgaaagaagcttttcccttcggtccggtccag 907  
 Db |||||  
 QY 944 ttccaagccttacataaactctccgattctccagcagcggaaactaaatcaccggagcg 1003  
 Db |||||  
 QY 908 tcccc---ccttctctaacttttccgtttcacaagaaaacggagaaaacagcggatg 964  
 Db |||||  
 QY 1004 tggagattcagc---ttgatccgcaacagatgatgcagcagatttccgattatcagagact 1060  
 Db |||||

Db 965 tcccgatcttctcaatgacaagcccgctggttcttccacgacttcgccatcaccaaaaact 1024  
 QY 1061 tcgtcgtcgtactgaccagcaagtcggttttcaagctgcggagatgattcccggtgggt 1120  
 Db |||||  
 Db 1025 acgcgattttcccgagatccaaatcggaatgagcccaatgagatgctgggtggggat 1084  
 QY 1121 ctccggtgggtttacgacaagaacaggtgcgaagatttcgggatttttagacaataacgccc 1180  
 Db |||||  
 Db 1085 ccccggtagcgcgagcgtggaaggtgctcgcggtcggttgcctcccgtaacgca 1144  
 QY 1181 aagatttcgaacacattaagtggattgattccagattgcttccagattgcttccatctcga 1240  
 Db |||||  
 Db 1145 aagacgagcccgagatgaagtggttccgggttccgggttttaattgatacatcatca 1204  
 QY 1241 acgcttgggaagagccgacaacacagatgaagtcgctcgtgtag 1282  
 Db |||||  
 Db 1205 atgcatggagagatggcggagatacgggtggtgatggtgg 1246

RESULT 14  
 AAC57157  
 ID AAC57157 standard; DNA; 491 BP.  
 XX  
 AC AAC57157;  
 XX  
 DT 25-JAN-2001 (first entry)  
 XX  
 DE Pinus radiata transcription factor DNA sequence #603.  
 KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
 KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
 KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
 KW homeodomain zipper; LIM domain; AP2; ERBs; zinc finger domain;  
 KW type 2 Cys2His2; CCAAT box element; MYB; ss.  
 XX  
 OS Pinus radiata.  
 XX  
 PN WO200053724-A2.  
 XX  
 PD 14-SEP-2000.  
 XX  
 PF 09-MAR-2000; 2000WO-US06112.  
 XX  
 PR 11-MAR-1999; 99US-0266513.  
 PR 18-AUG-1999; 99US-0149485.  
 XX  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
 XX  
 PI Wood M, McGrath A, Shenk MA, Glenn M;  
 XX WPI; 2000-579369/54.  
 XX  
 DR New isolated polynucleotide encoding a plant transcription factor for  
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
 PT having modified gene expression or modified activity of a polypeptide  
 PT  
 XX  
 PS Claim 1; Page 598; 747pp; English.  
 XX  
 CC The present invention relates to novel plant transcription factors from  
 CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
 CC sequence for one such transcription factor. The transcription factor may  
 CC be used to produce a plant having modified gene expression such as a  
 CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
 CC mahogany species or to modify the activity of a polypeptide in a plant.  
 CC The transcription factors of the present invention are members from the  
 CC following families of regulatory proteins: bZIP, bZIP family of G-box  
 CC binding factors, basic helix-loop-helix zipper,  
 CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
 CC and ERBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
 CC and MYB.  
 XX

SQ Sequence 491 BP; 86 A; 143 C; 171 G; 91 T; 0 other;

Query Match 8.3%; Score 149.4; DB 21; Length 491;  
Best Local Similarity 59.2%; Pred. No. 1.1e-37;  
Matches 255; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

Qy 436 cccgattccatcaaggaggtatgtgcgcaacgagagtaaccacttcacgagccggtg 495  
Dy 1 ccgagtgcttagatggagtttaagtcgcaatggcggaatccccgtttcaaacccgc 60  
Qy 496 acaggtccaccacttttcgagcgagacggtatggttcacgcccgtcaaattcgaacacggt 555  
Dy 61 ggcggccaccattttatgcagcgagatggaatgatacatgcgtagacgtgagacaggg 120  
Qy 556 tcagtagctacgttcggtttactcagactaacgggtttgttcagaaacgtcaattg 615  
Dy 121 aaggttagttacagttgcggtttcagcgagacggaaggctcgtagcgagagcgggcg 180  
Qy 616 ggtcgacgggtttccccaagccatcggtgagcttcacggccacacggtattgccga 675  
Dy 181 gggcgaggttttaccgaagccatcggtgcaactccacggccacggcggtggtgcgc 240  
Qy 676 ctcatgctattctacgcagagctgcagcgggtatagtcgacccgggcacacggacccgt 735  
Dy 241 ctgctgcgtatggtgcgggggctctgcgggtcggtcacaacgggaaggatgggc 300  
Qy 736 gtatgaacgcgggtttggtctatttcgaatggcggttattggtatgctcggaagatgat 795  
Dy 301 gtggttaatgcgggtggctcttttaacggcgctgcgtctatgtccgaagacgat 360  
Qy 796 ttaccttcaagttcagatcacctcccgaatggagatttaaaacggttgcgttcgat 855  
Dy 361 ctcccgatgcgtcagggtagcgggtgacggcgatctggtgacgacgggcaggttcgat 420  
Qy 856 ttgtaggaca 866  
Dy 421 ttcgacgggca 431

RESULT 15  
AAC57162  
ID AAC57162 standard; DNA; 386 BP.

XX AC AAC57162;

XX DT 25-JAN-2001 (first entry)

XX DE Pinus radiata transcription factor DNA sequence #608.

XX KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;  
KW type 2 Cys2His2; CCAAT box element; MYB; ss.

XX OS Pinus radiata.

XX FN WO200053724-A2.

XX PD 14-SEP-2000.

XX PF 09-MAR-2000; 2000WO-US06112.

XX PR 11-MAR-1999; 99US-0266513.

XX PR 18-AUG-1999; 99US-0149485.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.  
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX FI Wood M, McGrath A, Shenk MA, Glenn M;

XX DR WPI; 2000-579369/54.

XX

PT New isolated polynucleotide encoding a plant transcription factor for  
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
PT having modified gene expression or modified activity of a polypeptide

XX Claim 1; Page 599; 747pp; English.

XX CC The present invention relates to novel plant transcription factors from  
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
CC sequence for one such transcription factor. The transcription factor may  
CC be used to produce a plant having modified gene expression such as a  
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
CC mahogany species or to modify the activity of a polypeptide in a plant.  
CC The transcription factors of the present invention are members from the  
CC following families of regulatory proteins: bZIP, bZIP family of G-box  
CC binding factors, basic helix-loop-helix zipper,  
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
CC and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
CC and MYB.

XX SQ Sequence 386 BP; 89 A; 105 C; 106 G; 86 T; 0 other;

Query Match 8.0%; Score 143.4; DB 21; Length 386;  
Best Local Similarity 70.0%; Pred. No. 8.4e-36;  
Matches 208; Conservative 0; Mismatches 86; Indels 3; Gaps 1;

Qy 1103 agatgattccggtgggtctccggtggtttacgacaagaacaaggtcgcaagattcggga 1162

Dy 1 aaatgatcagaggcggtttccagttattcacacaagaagaaggtcccggtcgggc 60

Qy 1163 ttttagacaataacgcgaagatttcacgaacattaaagtggattgactccagattgct 1222

Dy 61 ttctgccaaaatgcttctgacgagagtgagctgaaatggatcgaggtcccgattgct 120

Qy 1223 tctgcttccatctctggaacgcttgggaagagccagaacagatgaagtcgtgatag 1282

Dy 121 tctgcttccatctctggaacgcttgggaaga---aggagaagacgaggttgcgtccatcg 177

Qy 1283 ggtcctgtagactccaccagactcaattttcaacgagctgacgagaatcttcaagagtg 1342

Dy 178 gctcctgtagaccgcccgagcgcattttcaacgaatctgacagcgcgtcgaggatg 237

Qy 1343 tctgtctgaaatccgctgaattcctcaaacccggtgaatcaactcgcccgatca 1399

Dy 238 ttctgctggaaattcggctcaatctcaaacccggttgcaccacgacgagatca 294

Search completed: July 24, 2002, 06:25:34

Job time: 6890 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2002, 03:54:00 ; Search time 57.66 Seconds  
(without alignments)  
7668.059 Million cell updates/sec

Title: US-09-758-269-5  
Perfect score: 1800  
Sequence: 1 atggctttcttcagcgcaac.....tggcgaagcaggctgtgtga 1800

Scoring table:  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38.8	2.2	7218	1 US-08-232-463-14	Sequence 14, Appl
2	34.8	1.9	4695	6 5225348-3	Patent No. 5225348
3	34.2	1.9	1221	1 US-08-434-881-1	Sequence 1, Appl
4	34.2	1.9	1221	3 US-08-977-771-1	Sequence 1, Appl
5	34.2	1.9	1221	4 US-09-361-773-1	Sequence 1, Appl
6	34.2	1.9	1370	4 US-09-026-408-12	Sequence 12, Appl
7	34.2	1.9	1371	4 US-09-026-408-1	Sequence 1, Appl
8	33.4	1.9	788	4 US-08-991-789A-177	Sequence 177, App
9	33.4	1.9	788	4 US-09-062-451-177	Sequence 177, App
10	33.4	1.9	1047	5 PCT-US96-10986-11	Sequence 11, Appl
11	33.4	1.9	1753	6 5225348-2	Sequence 11, Appl
12	33.4	1.9	3680	2 US-08-494-907-1	Patent No. 5225348
13	33.4	1.9	3680	5 PCT-US96-10986-1	Sequence 1, Appl
14	33.4	1.9	5076	2 US-08-494-907-2	Sequence 1, Appl
15	33.4	1.9	5076	5 PCT-US96-10986-2	Sequence 2, Appl
16	33.4	1.9	6170	2 US-08-494-907-4	Sequence 2, Appl
17	33.4	1.9	6170	5 PCT-US96-10986-4	Sequence 4, Appl
18	33.4	1.9	6387	2 US-08-494-907-3	Sequence 3, Appl
19	33.4	1.9	6387	5 PCT-US96-10986-3	Sequence 3, Appl
20	33.4	1.9	7198	4 US-08-994-035C-4	Sequence 4, Appl
21	33.4	1.9	4411529	4 US-09-103-840A-1	Sequence 1, Appl
22	33.4	1.9	607	4 US-09-385-982-525	Sequence 525, App
23	33	1.8	607	2 US-08-371-377-16	Sequence 16, Appl
24	33	1.8	2128	4 US-09-036-987A-1	Sequence 1, Appl
25	33	1.8	80161	3 US-09-036-987A-1	Sequence 1, Appl
26	33	1.8	80161	4 US-09-370-700-1	Sequence 1, Appl
27	32.6	1.8	1092	4 US-09-077-675A-15	Sequence 15, Appl

Sequence 14, Appl  
Sequence 14, Appl  
Sequence 33, Appl  
Sequence 33, Appl  
Sequence 33, Appl  
Sequence 33, Appl  
Sequence 33, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 25, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 448, App  
Sequence 14, Appl  
Sequence 12, Appl  
Sequence 12, Appl

ALIGNMENTS

RESULT 1  
US-08-232-463-14  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-F18  
US-08-232-463-14

```
Query Match      2.2%; Score 38.8; DB 1; Length 7218;
Best Local Similarity 7.9%; Pred. No. 0.049;
Matches 16; Conservative 112; Mismatches 74; Indels 0; Gaps 0;

Qy 21 ggctgagggtttctggagatgcttgggtggaatcatactcagcgccattatcgcttctc 80
Db 1041 GGCTGAGGTCAGGAGCTTGGCATYVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVV 1100

Qy 81 tcaagctccgacttgagttatttagctcttacctatgcccagtcgtgcacacgttaa 140
Db 1101 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1160

Qy 141 gctcaatgttcactgcgcttcacactccctcagctcttcattcccttaagcaatc 200
Db 1161 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1220

Qy 201 aaactctcccgcattgtgtt 222
Db 1221 YYYYYYYYYYYYYYYYYYYY 1242

RESULT 2
5225348-3
;PATENT NO. 5225348
;APPLICANT: HAGATA, SHIGEKAZU;SUGANO, SUMIO;KIM, DONG W.;
;UETSUKI, TAICHI;KAZIRO, YOSHITO
;TITLE OF INVENTION: DNA FRAGMENT AND EXPRESSION PLASMID
;CONTAINING THE DNA FRAGMENT
;NUMBER OF SEQUENCES: 9
;CURRENT APPLICATION DATA:
;APPLICATION NUMBER: US/07/447,823
;FILING DATE: 08-DEC-1989
;SEQ ID NO:3:
;LENGTH: 4695
5225348-3

Query Match      1.9%; Score 34.8; DB 6; Length 4695;
Best Local Similarity 48.1%; Pred. No. 0.78;
Matches 99; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

Qy 340 cctaaacggctgatccctcagtggtcagatcccgaaattttgctccggtgaatgaacag 399
Db 3386 cctgattggattggccacacggctcacattgcagcaagtttgcgtgagctgaaggaaag 3445

Qy 400 cccgtccggcgttaattcttcgggtggtcggaaacttccgattccatcaaaaggagtgat 459
Db 3446 attgatccgcttctggtaaaagctggaagatggcctaaattcttgagctcgtgat 3505

Qy 460 gtgcgcaacgagctaacccacttcacagcgggtgacaggtcaccactttctcgacgga 519
Db 3506 gctgcattgtgatgtgttctcgggaagcccatgtgtgtgagagcttctcagactat 3565

Qy 520 gacggtatggttcacgcgcgtcaaat 545
Db 3566 ccacctttgggtgaaggatgactactt 3591

RESULT 3
US-08-434-881-1/c
;Sequence 1, Application US/08434881
;Patent No. 5804376
;GENERAL INFORMATION:
;APPLICANT: Braxton, Scott M.
;APPLICANT: Wilde, Craig G.
;APPLICANT: Diep, Dinh
;TITLE OF INVENTION: Pancreas-Derived Serpin
;NUMBER OF SEQUENCES: 2
;CORRESPONDENCE ADDRESS:
;ADDRESSEE: Incyte Pharmaceuticals, Inc.
;STREET: 3330 Hillview Avenue
;CITY: Palo Alto
;STATE: California
```

```
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,881
; FILING DATE: Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Luther, Barbara J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PF0035 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-852-0195
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1221 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; LIBRARY: Pancreas
; CLONE: 222689
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1221
; US-08-434-881-1

Query Match      1.9%; Score 34.2; DB 1; Length 1221;
Best Local Similarity 54.3%; Pred. No. 0.51;
Matches 69; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 97 agttattgtagctcttaactatggccagtcggtgcacgtaagctcaatgtttctatct 156
Db 788 ACTTCTTCTATATCCATACCTTCTTCGAGGAAGTATGATAATTAGCTAAATTCATCACC 729

Qy 157 gcgcttcacactctccagctcttcttccctaaagcaatcatcaactctccgcccatt 216
Db 728 TTGTAAGACAACTTAAACTTGGTAGTTAGGGAAGATTGAGAAATATACCATATTTT 669

Qy 217 gttgta 223
Db 668 GTTCTCA 662

RESULT 4
US-08-977-771-1/c
;Sequence 1, Application US/08977771
;Patent No. 6013448
;GENERAL INFORMATION:
;APPLICANT: Braxton, Scott M.
;APPLICANT: Wilde, Craig G.
;APPLICANT: Diep, Dinh
;TITLE OF INVENTION: Pancreas-Derived Serpin
;NUMBER OF SEQUENCES: 2
;CORRESPONDENCE ADDRESS:
;ADDRESSEE: Incyte Pharmaceuticals, Inc.
;STREET: 3330 Hillview Avenue
;CITY: Palo Alto
;STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```

RESULT      5
US-09-361-773-1/C
; Sequence 1, Application US/09361773
; Patent No. 6197519
; GENERAL INFORMATION:
; APPLICANT: Braxton, Scott M.
; APPLICANT: Wilde, Craig G.
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Pancreas-Derived Serpin
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3330 Hillview Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/361,773
; FILING DATE:
; CLASSIFICATION:

```

RESULT 6  
US-09-026-408-12/c  
; Sequence 12, Application US/09026408  
; Patent No. 630338  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: PANCREAS-DERIVED PLASMINOGEN ACTIVATOR  
; TITLE OF INVENTION: INHIBITOR  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVENUE, SUITE 600  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/026.408  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; Prior Application Data:  
; APPLICATION NUMBER: US 08/934,011  
; FILING DATE: 15-AUG-1997





CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Potter, Jane E. R.  
REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 210121.419C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 177:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 788 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 177:  
US-08-991-789A-177

Query Match 1.9%; Score 33.4; DB 4; Length 788;  
Best Local Similarity 48.2%; Pred. No. 0.71;  
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;  
QY 340 cctaaacggctgacctagttcagatcgcggaattttgctccggtgaatgaacag 399  
DB 536 CTGTATTGGATTGCCACAGCTACATTGCTGAGCTGAAGGAAAG 595  
QY 400 cccgtccggcgtaattcttcggtggtcggaatactcccgattccatcaaaaggagtgtat 459  
DB 596 ATTGATCGCGTTCGTGTAAGGCTGGAAGATGGCCCTAAATTTCTTGAAGTCTGGTGTAT 655  
QY 460 gtgcgaacggagctaacccacttcacgagcgggtgacaggtcaccactttcttcacgga 519  
DB 656 GCTGCAATTTGATATGTTCTTCTGCAAGCCCATGTGTGTTGAGAGCTTCTCAGACTAT 715  
QY 520 gacggtatggttcac 534  
DB 716 CCACCTTTGGGTGCG 730

RESULT 9  
US-09-062-451-177  
; Sequence 177, Application US/09062451  
; Patent No. 6344550  
; GENERAL INFORMATION:  
; APPLICANT: Frudakis, Tony N.  
; APPLICANT: Smith, John M.  
; APPLICANT: Reed, Steven G.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER  
; NUMBER OF SEQUENCES: 297  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/062,451  
; FILING DATE: 04-APR-1997  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.419C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 177:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 788 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-062-451-177

Query Match 1.9%; Score 33.4; DB 4; Length 788;  
Best Local Similarity 48.2%; Pred. No. 0.71;  
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;  
QY 340 cctaaacggctgacctagttcagatcgcggaattttgctccggtgaatgaacag 399  
DB 536 CTGTATTGGATTGCCACAGCTACATTGCTGAGCTGAAGGAAAG 595  
QY 400 cccgtccggcgtaattcttcggtggtcggaatactcccgattccatcaaaaggagtgtat 459  
DB 596 ATTGATCGCGTTCGTGTAAGGCTGGAAGATGGCCCTAAATTTCTTGAAGTCTGGTGTAT 655  
QY 460 gtgcgaacggagctaacccacttcacgagcgggtgacaggtcaccactttcttcacgga 519  
DB 656 GCTGCAATTTGATATGTTCTTCTGCAAGCCCATGTGTGTTGAGAGCTTCTCAGACTAT 715  
QY 520 gacggtatggttcac 534  
DB 716 CCACCTTTGGGTGCG 730

RESULT 10  
US-08-494-907-11/c  
; Sequence 11, Application US/08494907  
; Patent No. 5955298  
; GENERAL INFORMATION:  
; APPLICANT: Thomasow, Linda S  
; APPLICANT: Bangera, Mahalaxmi  
; APPLICANT: Weller, David M  
; APPLICANT: Cook, R. James  
; TITLE OF INVENTION: Sequences for Production of  
; TITLE OF INVENTION: 2,4-diacetylphloroglucinol and Methods  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Margaret A. Connor, USDA-ARS  
; STREET: 800 Buchanan Street  
; CITY: Albany  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94710  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/494,907  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connor, Margaret A  
; REGISTRATION NUMBER: 30043  
; REFERENCE/DOCKET NUMBER: 0009.95  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 559-6067  
; TELEFAX: (510) 559-5777  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1047 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO

```
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pseudomonas fluorescens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1..1047)
; OTHER INFORMATION: /note= "ph1D DNA sequence. SEQ ID
; US-08-494-907-11
; OTHER INFORMATION: NO:12 is translation (protein) of SEQ ID NO:11."
;
Query Match      1.9%; Score 33.4; DB 2; Length 1047;
Best Local Similarity 48.2%; Pred. No. 0.85;
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 585 gactaacgggtttgttcagggaacgtcaattg99gtgacgggtttttcccccagaagccatcgg 644
Db 937 GGCTAAACGGATGATTCAAAACACCCCAAGTCAATGAACGGTATCTGGTCTTGCCCATCGA 878

Qy 645 tgagcttcacggccacacacggattgcccgcactcatgctattctacgccagagctgcagc 704
Db 877 TGAACCTTCGGTGCATACCGGCTTTACCCACCGCAGCATCGTGTATGAGCGAGAGGCTCG 818

Qy 705 cggatatgtagcaccggccacacggacgggtgtagctaacgccgggtttggtctatttcaa 764
Db 817 CCGCATGCTTCAATCGCGCGCGCCAGGCCATTGAAAACGCGGGCTTGACCCAGGACGA 758

Qy 765 tggccgggtattggc 779
Db 757 CATCCGGATGGTGC 743

RESULT 11
PCT-US96-10986-11/c
; Sequence 11, Application PC/TUS9610986
; GENERAL INFORMATION:
; TITLE OF INVENTION: Sequences for Production of
; 2,4-Diacetylphloroglucinol and Methods
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stephan A. Pendorf, DOMINIK & STEIN
; STREET: 600 N. West Shore Boulevard, Suite 1000
; CITY: Tampa
; STATE: FL
; COUNTRY: USA
; ZIP: 33609
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10986
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pendorf, Stephan A.
; REGISTRATION NUMBER: 32665
; REFERENCE/DOCKET NUMBER: A700.320
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (813) 289-2966
; TELEFAX: (813) 289-2967
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1047 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pseudomonas fluorescens
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1..1047)
; OTHER INFORMATION: /note= "ph1D DNA sequence. SEQ ID
; PCT-US96-10986-11
; OTHER INFORMATION: NO:12 is translation (protein) of SEQ ID NO:11."
;
Query Match      1.9%; Score 33.4; DB 5; Length 1047;
Best Local Similarity 48.2%; Pred. No. 0.85;
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 585 gactaacgggtttgttcagggaacgtcaattg99gtcgcacgggtttttcccccagaagccatcgg 644
Db 937 GGCTAAACGGATGATTCAAAACACCCCAAGTCAATGAACGGTATCTGGTCTTGCCCATCGA 878

Qy 645 tgagcttcacggccacacacggattgcccgcactcatgctattctacgccagagctgcagc 704
Db 877 TGAACCTTCGGTGCATACCGGCTTTACCCACCGCAGCATCGTGTATGAGCGAGAGGCTCG 818

Qy 705 cggatatgtagcaccggccacacggacgggtgtagctaacgccgggtttggtctatttcaa 764
Db 817 CCGCATGCTTCAATCGCGCGCGCCAGGCCATTGAAAACGCGGGCTTGACCCAGGACGA 758

Qy 765 tggccgggtattggc 779
Db 757 CATCCGGATGGTGC 743

RESULT 12
5225348-2
; Patent No. 5225348
; APPLICANT: HAGATA, SHIGEKAZU; SUGANO, SUMIO; KIM, DONG W.;
; UETSUKI, TAICHI; KAZIRO, YOSHITO
; TITLE OF INVENTION: DNA FRAGMENT AND EXPRESSION PLASMID
; CONTAINING THE DNA FRAGMENT
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/447,823
; FILING DATE: 08-DEC-1989
; SEQ ID NO:2:
; LENGTH: 1753
5225348-2

Query Match      1.9%; Score 33.4; DB 6; Length 1753;
Best Local Similarity 48.2%; Pred. No. 1.2;
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 340 cctaaacggctgacctagttctcagatcgccggaaattttgctcgggtgaatgaacag 399
Db 1137 cctgattggattgccacacggctccatcgcatgcaagtttgcgagcgaaggaagaaag 1196

Qy 400 cccgtccgggtaattcttcgggtggtcggaacacttcccgaattccatcaaggaggtgat 459
Db 1197 attgatcgccgttctgtgtataaaagctggaagatggccctaaattcttgagctcgtgat 1256

Qy 460 gtgcgaacggagctaacccacttcacgacggcggtgcaggtcacccacttttccgacgga 519
Db 1257 gctgccattgtgatgtgttcttcgggaagcccatgtgtgtgagagcttctcagactat 1316

Qy 520 gacggtatggttcac 534
Db 1317 ccacctttgggtcgc 1331

RESULT 13
US-08-494-907-1/c
; Sequence 1, Application US/08494907
; Patent No. 5955298
; GENERAL INFORMATION:
; APPLICANT: Thomasow, Linda S
; APPLICANT: Bangera, Mahalaxmi
```

```

; APPLICANT: Weller, David M
; APPLICANT: Cook, R. James
; TITLE OF INVENTION: Sequences for Production of
; TITLE OF INVENTION: 2,4-Diacetylphloroglucinol and Methods
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Margaret A. Connor, USDA-ARS
; STREET: 800 Buchanan Street
; CITY: Albany
; STATE: CA
; COUNTRY: USA
; ZIP: 94710
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/494,907
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Connor, Margaret A
; REGISTRATION NUMBER: 30043
; REFERENCE/DOCKET NUMBER: 0009.95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 559-6067
; TELEFAX: (510) 559-5777
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: Pseudomonas fluorescens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: complement (2889..3680)
; OTHER INFORMATION: /note="phlA, transcribed from
; OTHER INFORMATION: right to left"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: complement (1210..1917)
; OTHER INFORMATION: /note="phlB, transcribed from
; OTHER INFORMATION: right to left"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: complement (1689..2855)
; OTHER INFORMATION: /note="phlC, transcribed from
; OTHER INFORMATION: right to left"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: complement (2..1048)
; OTHER INFORMATION: /note="phlD, transcribed from
; OTHER INFORMATION: right to left"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 2118..3371
; OTHER INFORMATION: /note="phlR, transcribed from left
; OTHER INFORMATION: to right"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..3680
; OTHER INFORMATION: /note="SEQ ID NO:1 contains genes
; OTHER INFORMATION: necessary for Phl synthesis."
US-08-494-907-1

```

Query Match 1.9%; Score 33.4; DB 2; Length 3680;

```

Best Local Similarity 48.2%; Pred. No. 1.9;
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 585 gactaacgggtttgttcaggaacgtcaattgggtcgaccgggtttccccaagccatcgg 644
DB 938 GGCTAAACGGATGATTCAAAACACCACTCAATGAACGGTATCTGGTCTTGCCCATCGA 879
QY 645 tgagcttcacggccacacgggtattgcccagactatgtattctacgccagagctgcagc 704
DB 878 TGAACCTGGGTGCATACCGGCTTTACCCACCGCAGCATCGTGTATGACGAGAGGCTCG 819
QY 705 cggatatgtcagccggcacacgggaacgggtgtagctaacgccgggttgggtattttaa 764
DB 818 CCGCATGCTTCAATCGCGCGCGCCGCGCCATTGAAACGCGGCTTGACCCAGGACCGA 759
QY 765 tggccgggtattggc 779
DB 758 CATCCGGATGGTCG 744

RESULT 14
PCT-US96-10986-1/c
; Sequence 1, Application PC/TUS9610986
; GENERAL INFORMATION:
; TITLE OF INVENTION: Sequences for Production of
; TITLE OF INVENTION: 2,4-Diacetylphloroglucinol and Methods
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stephan A. Pendorf, DOMINIK & STEIN
; STREET: 600 N. West Shore Boulevard, Suite 1000
; CITY: Tampa
; STATE: FL
; COUNTRY: USA
; ZIP: 33609
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10986
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pendorf, Stephan A.
; REGISTRATION NUMBER: 32665
; REFERENCE/DOCKET NUMBER: A700.320
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (813) 289-2966
; TELEFAX: (813) 289-2967
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: Pseudomonas fluorescens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: complement (2889..3680)
; OTHER INFORMATION: /note="phlA, transcribed from
; OTHER INFORMATION: right to left"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: complement (1210..1917)
; OTHER INFORMATION: /note="phlB, transcribed from
; OTHER INFORMATION: right to left"
; FEATURE:
; NAME/KEY: misc feature

```

LOCATION: complement (1689..2855)  
OTHER INFORMATION: /note= "phlC, transcribed from  
OTHER INFORMATION: right to left"

## FEATURE:

NAME/KEY: misc feature  
LOCATION: complement (2..1048)  
OTHER INFORMATION: /note= "phlD, transcribed from  
OTHER INFORMATION: right to left"

## FEATURE:

NAME/KEY: misc feature  
LOCATION: 2118..3371  
OTHER INFORMATION: /note= "phlR, transcribed from left  
OTHER INFORMATION: to right"

## FEATURE:

NAME/KEY: misc feature  
LOCATION: 1..3680  
OTHER INFORMATION: /note= "SEQ ID NO:1 contains genes  
OTHER INFORMATION: necessary for Phl synthesis."

PCT-US96-10986-1

Query Match 1.9%; Score 33.4; DB 5; Length 3680;  
Best Local Similarity 48.2%; Pred. No. 1.9;  
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 585 gactaacccggtttgttcaggaaacgtcaattgggtcgaccggttttccccaaagccatcgg 644

Db 938 GGCATAACGGATGATTCATAAACACACCCAAAGTCAATGAACGGTATCTGGTTCGCCCATCGA 879

Qy 645 tgagcttcacggccacacccggtattggccgactcatgtattctacgcagagctgcagc 704

Db 878 TGAATTCCGCTGCATACCGGCTTTACCCACCGCAGCATCTGTATGACGAGAGGCTCG 819

Qy 705 cggctatagtcgaccggccacacgggaacgggtgtagtaaacgccgggtttgtctatttcaa 764

Db 818 CCGCATGTCTTCAATCGCGCGCGCCAGCCATTGAAAACCGGGCTTGACCCAGGACGA 759

Qy 765 tggccggttatggc 779

Db 758 CATCCGGATGGTCGC 744

## RESULT 15

US-08-494-907-2/c

Sequence 2, Application US/08494907

Patent No. 5955298

GENERAL INFORMATION:

APPLICANT: Thomasow, Linda S

APPLICANT: Bangera, Mahalaxmi

APPLICANT: Weller, David M

APPLICANT: Cook, R. James

TITLE OF INVENTION: Sequences for Production of

TITLE OF INVENTION: 2,4-Diacetylphloroglucinol and Methods

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Margaret A. Connor, USDA-ARS

STREET: 800 Buchanan Street

CITY: Albany

STATE: CA

COUNTRY: USA

ZIP: 94710

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/494,907

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Connor, Margaret A

REGISTRATION NUMBER: 30043

REFERENCE/DOCKET NUMBER: 0009.95  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (510) 559-6067

TELEFAX: (510) 559-5777

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 5076 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Pseudomonas fluorescens

FEATURE:

NAME/KEY: misc feature

LOCATION: complement (4285..5076)

OTHER INFORMATION: /note= "phlA, transcribed from

OTHER INFORMATION: right to left"

FEATURE:

NAME/KEY: misc feature

LOCATION: complement (2606..3313)

OTHER INFORMATION: /note= "phlB, transcribed from

OTHER INFORMATION: right to left"

FEATURE:

NAME/KEY: misc feature

LOCATION: complement (3085..4251)

OTHER INFORMATION: /note= "phlC, transcribed from

OTHER INFORMATION: right to left"

FEATURE:

NAME/KEY: misc feature

LOCATION: complement (1398..2444)

OTHER INFORMATION: /note= "phlD, transcribed from

OTHER INFORMATION: right to left"

FEATURE:

NAME/KEY: misc feature

LOCATION: 3514..4767

OTHER INFORMATION: /note= "phlR, transcribed from left

OTHER INFORMATION: to right"

FEATURE:

NAME/KEY: misc feature

LOCATION: complement (2..1270)

OTHER INFORMATION: /note= "phlE, transcribed from

OTHER INFORMATION: right to left"

FEATURE:

NAME/KEY: misc feature

LOCATION: 1..5076

OTHER INFORMATION: /note= "SEQ ID NO:2 contains genes

OTHER INFORMATION: involved in synthesis, activity, and/or export of

OTHER INFORMATION: Phl."

US-08-494-907-2

Query Match

Best Local Similarity 1.9%; Score 33.4; DB 2; Length 5076;

Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 585 gactaacccggtttgttcaggaaacgtcaattgggtcgaccggttttccccaaagccatcgg 644

Db 2334 GGCATAACGGATGATTCATAAACACACCCAAAGTCAATGAACGGTATCTGGTTCGCCCATCGA 2275

Qy 645 tgagcttcacggccacacccggtattggccgactcatgtattctacgccagagctgcagc 704

Db 2274 TGAACCTTCGGTGCATACCGGCTTTACCCACCGCAGCATCTGTATGACGAGAGGCTCG 2215

Qy 705 cggctatagtcgaccggccacacgggaacgggtgtagtaaacgccgggtttgtctatttcaa 764

Db 2214 CCGCATGTCTTCAATCGCGCGCGCCAGGCCATTGAAAACCGGGCTTGACCCAGGACGA 2155

Qy 765 tggccggttatggc 779

Db 2154 CATCCGGATGGTCGC 2140

Search completed: July 24, 2002, 06:21:27  
Job time: 8847 sec

---

**This Page Blank (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2002, 05:14:15 ; Search time 1694.89 Seconds  
(without alignments)  
14333.980 Million cell updates/sec

Title: US-09-758-269-5  
Perfect score: 1800  
Sequence: 1 atggctcttcacggcaac.....tggcgaagcaggtcgtgtga 1800

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues  
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	425.6	23.6	720	10	BM412731
2	417	23.2	805	10	BM408615
3	416.2	23.1	781	12	BH549344
4	382.8	21.3	643	9	AW933324
5	375.6	20.9	787	10	BM408565
6	344	19.1	592	10	BF113346
7	332.4	18.5	566	10	BE461924
8	323.2	18.0	553	10	BM085672
9	318	17.7	559	10	BM085005
10	316.6	17.6	520	10	BE432853
11	313.8	17.4	742	12	BH559796
12	312.8	17.4	594	12	BH458011
13	309.4	17.2	547	10	BI974879
14	308.6	17.1	547	10	BE437072
15	303.4	16.9	617	10	BE458861
16	300.4	16.7	546	10	BM084948
17	295.8	16.4	618	9	AW443298

18	294	16.3	564	10	BM085488
19	292.8	16.3	627	10	BF051297
20	289.2	16.1	509	10	BE451573
21	288.8	16.0	495	10	BE434930
22	276.2	15.3	490	10	BF050563
23	275.4	15.3	657	9	AA556214
24	267.4	14.9	503	12	B27476
25	254.8	14.2	566	10	BE459895
26	249.2	13.8	450	9	AW930245
27	248	13.8	552	10	BG039692
28	245	13.6	778	12	BH579654
29	235.8	13.2	600	12	BH458018
30	233.2	13.0	424	12	BH496327
31	224.4	12.5	337	12	BH579641
32	220.4	12.2	410	10	BE522530
33	217	12.1	446	10	BE471089
34	216.6	12.0	713	10	BI928174
35	209.6	11.6	495	9	AU084510
36	207	11.5	502	12	CNS000KV
37	205.6	11.4	590	10	BG592987
38	203.2	11.3	447	9	AW933245
39	202.8	11.3	492	9	AU084760
40	201.4	11.2	412	9	AV419581
41	195.2	10.8	680	12	AQ969739
42	185.8	10.3	549	10	BF425834
43	173.6	9.6	500	9	AW289745
44	164.6	9.1	678	12	BH593312
45	152.6	8.5	267	10	BE529289

ALIGNMENTS

RESULT 1  
BM412731  
LOCUS BM412731 720 bp mRNA linear EST 22-JAN-2002  
DEFINITION EST587047 tomato breaker fruit Lycopersicon esculentum cDNA clone  
ACCESSION CLE60N24 5' end, mRNA sequence.  
VERSION BM412731 GI:18264350  
KEYWORDS EST.  
SOURCE tomato.  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.  
1 (bases 1 to 720)  
REFERENCE Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A., Tsai  
AUTHORS J., Bougri,O., Kirkness,E., Utterback,T., Van Aken,S., Ronning  
,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.  
Generation of ESTs from tomato fruit tissue, breaker stage (2002)  
Unpublished (2002)  
Contact: CUGI  
Clemson University  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics  
Institute  
Seq primer: T3.  
Location/Qualifiers  
1. 720  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CLE60N24"  
/clone\_lib="tomato breaker fruit"  
/tissue\_type="Pericarp"  
/dev\_stage="breaker"  
/lab\_host="SOUR"

FEATURES  
source  
1. 720  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CLE60N24"  
/clone\_lib="tomato breaker fruit"  
/tissue\_type="Pericarp"  
/dev\_stage="breaker"  
/lab\_host="SOUR"

Site\_1: EcoRI;  
Site\_2: XhoI; supplier: Boyce Thompson Institute;

sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

```
BASE COUNT      226 a 143 c 155 g 196 t
ORIGIN

Query Match      23.6%; Score 425.6; DB 10; Length 720;
Best Local Similarity 74.4%; Pred. No. 6.3e-116;
Matches 536; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

Qy 740 ctaacgcgggttggtctatttcaatggccgggtattggtctatgctcgaggatatttac 799
Db 1 CAAGCCGGGTTAGTCTATTTCATAAACCAGTATTACTTGTCTGAAGATGATTTCG 60

Qy 800 cttaccaagttcagatcactcccaatggagattttaaaaccggttggtcggttcgattttg 859
Db 61 CTTACCATGTAAGTAACACCCCGCGGATCTTTAAACAGAGGGTGCATTCGATTTCG 120

Qy 860 atggacaattagaatccacaatgattgcccaccggaaagtgcagccgggaatccgggtgaac 919
Db 121 ACGGCCAGCTAAATCCACCATGATAGCTCACCCAAAGCTCGACCCAGTTTCCGGTGAGC 180

Qy 920 tcttcgcttaagctacgacgtcgcttcaagccttactctaaataacttccgattctcac 979
Db 181 TATTGCTCTTAGCTACGATGATGATTCAGAACCCATACCTCAAGTACTTCAGATTTTCAA 240

Qy 980 cggacggaactaaatcccgacgctcgagatcagcttgatcagccaacgatgatgcacg 1039
Db 241 AAATGGGAAATCAATGATGTTGAAATCCAGTTGAGACCCCAACATGATGTCATG 300

Qy 1040 atttcgcgattacagagaacttcgctcgctacgtacgtgaccagcagagtcgttttcaagctgc 1099
Db 301 ATTTCCCAATTACTGAGAACTTCGTCGTCATCTCCTGTATCAACAAGTCGTTTTCAAGATGT 360

Qy 1100 cggagatgacccgggtgggtctccgggtggtttacgacaagaadaaggtcgcaagattcg 1159
Db 361 CTGAATGATCGTGAGGTTTACCCGGTGGTTTACGACAAGAACAAAGTTTCCCGATTG 420

Qy 1160 ggattttagacaaatcacgcgaagattcatcgaaacattaaagtggattgatgctccagatt 1219
Db 421 GTATTCTGATAAGTACCGGAAAGATGGGTCTGATTTGAAATGGTTGAGTACCTGATT 480

Qy 1220 gcttctgtccatctctggaacgcttgggaagaccggaacagatgaagtgcgtgga 1279
Db 481 GTTCTGTGTTCCACCTCTGGAATGCTTGGGAAGAGCAGAAACAGATGAAATCGTTGTAA 540

Qy 1280 tagggctctgatgactccaccagactcaattttcaacgagctctgacgagaaatctcaaga 1339
Db 541 TTGGTTCATGTATGACACACACAGACTCATTTTCAATGAAATGTGATGAAGGGCTAAAGA 600

Qy 1340 gtgtcctgtctgaaatccgctgaaatcctcaaaaccggtgaaatcaactccgctccgactca 1399
Db 601 GTGTTTATCCGAATTCGCTCTCAATTTTGAACAGGGAAATTCACAGAAATCCATAA 660

Qy 1400 tctccaacgaagatcaacaagtcaacctcgaagcagggtggttcaacgaacaatgctgcg 1459
Db 661 TCGAAACCCCGATGAACAAGTGGATTTAGAAAGCTGGAATGCTGAACCGAAACAAACTCG 720

RESULT 2
BM408615
LOCUS      BM408615
DEFINITION EST582942 tomato breaker fruit Lycopersicon esculentum cDNA clone
            cLEG45023 5' end, mRNA sequence.
ACCESSION  BM408615
VERSION    BM408615.1 GI:18260245
KEYWORDS   EST.
SOURCE     tomato.
ORGANISM   Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 805)  
Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A., Tsai,J., Bougri,O., Kirkness,E., Utterback,T., Van Aken,S., Rinning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.  
Generation of ESTs from tomato fruit tissue, breaker stage (2002) Unpublished (2002)

TITLE JOURNAL  
COMMENT  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics Institute

Seq primer: T3.

FEATURES  
source  
1. .805  
/location=Qualifiers  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLEG45023"  
/clone\_lib="tomato breaker fruit"  
/tissue\_type="Pericarp"  
/dev\_stage="breaker"  
/lab\_host="SOLR"  
/notes="Vector: pBluescriptSKmCUadapt; Site\_1: EcoRI; Site\_2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."  
262 a 144 c 180 g 219 t

BASE COUNT  
ORIGIN

Query Match 23.2%; Score 417; DB 10; Length 805;  
Best Local Similarity 71.8%; Pred. No. 2.5e-113;  
Matches 575; Conservative 0; Mismatches 220; Indels 6; Gaps 2;

```
Qy 899 tcgaccggaatccggtgaactcttcgctttaagctacgacgctcggttccaagccttacc 958
Db 6 TCGACCCAGTTTCCGGTGAGCTATTGCTCTTAGCTAGCTGATTCAGAAAGCCATACC 65

Qy 959 taaatacttcgattctcaccggacggaactaaatcacggacgctcgagattcagcttg 1018
Db 66 TCAAGTACTTCAGATTTTCAAAAAATGGGAAAAATCAATGATGTTGAAATTCAGTTG 125

Qy 1019 atcggccaacgatgatgcagatttcgcatcacgaaacttcgctcgctgaactgacc 1078
Db 126 AAGACCCAAACAAATGATGATTCGCAATTACTGAGAACTTCGTCGTCATTCCTGATC 185

Qy 1079 agcaagtcggtttcaagctgcggagatgatccgctgggtctccggtggtttacgaca 1138
Db 186 AACAAAGTCGTTTTCAAGATGTCGAAATGATCCGTGGAGGTTACCCGGTGGTTTACGACA 245

Qy 1139 agaacaaggtcgcaagattccgggatttttagaacaatacgcgcaagattcatcgaaacatta 1198
Db 246 AGAACAAAGTTTCCGATTTTGGTATTCTGGATAAGTAGCCGAAAGATGGGTCTGATTGA 305

Qy 1199 agtggattgatgctccagattgcttcctgcttcctccatctctggaacgcttgggaagaccag 1258
Db 306 AATGGGTTGAAGTACCTGATGTTTCTGTTTCCACCTCTGGAATGCTTGGGAAGAAGCAG 365

Qy 1259 aaacagatgaagtgcgtgcgtgaggtccctgatgactccaccagactcaattttcgaacg 1318
Db 366 AAACAGATGAAATCGTTGTAATCGTTTCATGATATGACACCCACAGACTCCATTTTCAATG 425

Qy 1319 agtctgacgagaatctcaagagtgctcctgtctgaaatccgctgaatctcaaaacccggtg 1378
Db 426 AATGTGATGAAGGGCTAAAGAGTGTGTTTTATCCGAAATCCGTCCTCAATTTTGAACACAGGA 485
```



QY 1379 aatcaactgcgtccgagatcttccaaagaaagatcaacaagtcaacccctgaagcggga 1438  
 |||||  
 Db 486 AATCAACAAGAAATCCATAATCGAAACCCCGGATGAACAAGTGAATTTAGAAAGCTGGAA 545  
 |||||  
 QY 1439 tgggtcaacagaaacatgcctcgccggtgaaaccaaattcgcttacttggcttagccgagc 1498  
 |||||  
 Db 546 TGGTGAACCCGAAACAACCTCGGAAGGAACACAGAGTATGCTTATTTGGCTATCGCTGAAC 605  
 |||||  
 QY 1499 cgtggcctaaagtctcaggattcgctaaagtgtatctcactactggagaagttaaagaac 1558  
 |||||  
 Db 606 CATGGCCAAAGTTCTGGTTTTCGAAAGTAAACCTGTTCAACGGTGAAGTTGAGAAAT 665  
 |||||  
 QY 1559 atctttacggcgataacccgttacggaggagcctctgtttctcccgagagaaggagga- 1617  
 |||||  
 Db 666 TCATTATTTGTTGACAAACAATATGTTGGGAACCTCTTTTTTACAGAGAAACCCACAG 725  
 |||||  
 QY 1618 ---gagggaagcagagatcacctctgtttcttcacgacgagaaacatggaatc 1673  
 |||||  
 Db 726 CAGGAAGAGACGATGGTTATATTTTACGTTTCAGATGAGAA-AGATGGGAATC 784  
 |||||  
 QY 1674 ggaattacagatagtttaacgc 1694  
 |||||  
 Db 785 AAAATGCAAAATGGTAACGC 805  
 |||||

RESULT 3  
 LOCUS BH549344 781 bp DNA linear GSS 14-DEC-2001  
 DEFINITION BOGRQ53TF BOGR Brassica oleracea genomic clone BOGRQ53, DNA sequence.  
 ACCESSION BH549344  
 VERSION BH549344.1 GI:17801124  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea.  
 ORGANISM Brassica oleracea.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 781)  
 Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.  
 Whole genome shotgun sequencing of Brassica oleracea  
 Unpublished (2001)  
 Other GSSs: BOGRQ53TR  
 Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TF  
 Class: sheared ends.

FEATURES  
 source  
 1..781  
 /organism="Brassica oleracea"  
 /strain="T01000D3"  
 /db\_xref="taxon:3712"  
 /clone="BOGRQ53"  
 /clone\_lib="BOGR"  
 /note="Vector: pHS01; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHS01 using BstXI linkers"  
 BASE COUNT 204 a 203 c 192 g 182 t  
 ORIGIN  
 Query Match 23.1%; Score 416.2; DB 12; Length 781;  
 Best Local Similarity 70.8%; Pred. No. 4.3e-113;  
 Matches 553; Conservative 0; Mismatches 228; Indels 0; Gaps 0;  
 QY 274 gggcgggggcagcgttgagcggcgagggttctctgtcagcagcagagagtaac 333  
 |||||  
 Db 1 GCAGCGGGGATTCGGATCGATGGCTGAGCGAGCGTTGATCTCAGTGAAGATACT 60  
 |||||

QY 334 ccgcttcctaaacggcgctgattcctagttcagatcgccgggaaattttgctccgggtgaat 393  
 |||||  
 Db 61 CCTCTTCCCAAAACCGCTGATCCAGTATTCAAATCGCCGGGAATTTATTTCCGGTGGCG 120  
 |||||  
 QY 394 gaacggccgctccggcgtaattcttcgggtggtcggaanaacttcccgattcccatcaagga 453  
 |||||  
 Db 121 GAATTTCCCGTCCGAGAGTGTCTCGACGTTGAAGGAACAATCCCTGACTGCATTAACGGA 180  
 |||||  
 QY 454 gfgtatgfcgcaacgggagctaacccacttcacgagcgggtgacaggtcaccactcttc 513  
 |||||  
 Db 181 GTTATGTCCTGTAACGGTGCATAATCGATGTTTCGAGCCCAATCGTGGGCATCATTTGTTTC 240  
 |||||  
 QY 514 gacggagacgggtatggttcacgcgcgtcgaattcgaacacgggttcagctagctacgcttgc 573  
 |||||  
 Db 241 GACGGAGACGGAATGGTTCACGCAGTTAAATTAACAACGGTTTACGAAGTACGCATGC 300  
 |||||  
 QY 574 cggtttactcagataaacgggttttccaaggaacgtcaattgggtcgacgggttttccccc 633  
 |||||  
 Db 301 CGGTTTACTTAAACCGAGAGATTAATTCAGAAACAAACGATTGGGTTCGACCGGTTTTCCG 360  
 |||||  
 QY 634 aaagccatcggtgagcttcacggccacacccggtattgccgactcatgctattctacgcc 693  
 |||||  
 Db 361 AAAGCAATCGCGAGCTTCACGCCACTCGGAATTCGCACGTCTGATGCTTTTTTACGGA 420  
 |||||  
 QY 694 agagctgcagccggtatagtcgacccgggcacacggggtgttagctaacgccggtttg 753  
 |||||  
 Db 421 CGTGGGCTTTGGGCTCGTCAACAATCAAAACGGGTCGGGGTAGCGAACCCGGTTTG 480  
 |||||  
 QY 754 gtcatttcaatggcggttattggctatgttggaggatgatttaccctaccgaagtcag 813  
 |||||  
 Db 481 GTTTACTTCAATAACCGGCTTTTAGCAATGTGAGAAGACGATTTACCGTACCAATTAATA 540  
 |||||  
 QY 814 atcactcccaatggagatttaaaaaccgttggtcgggtcgattttgatggacaattagaa 873  
 |||||  
 Db 541 ATAACTCAAAACCGGACCTCTCAAAACCGTTGGCGTTACGATTCGACGGTCAGTTAAAA 600  
 |||||  
 QY 874 tcacaatgattgcccccccgaaagtgcacccgggaatccgggtgaactcttcgctttaagc 933  
 |||||  
 Db 601 TCATCGATGATCGCCACCCCGAAACTCGACCCGGTTACAAAGGAGCTCCACCGCTAACG 660  
 |||||  
 QY 934 tacgagctcggtttcaagcccttaccctaaataacttccgatttcccggaagaaactaaa 993  
 |||||  
 Db 661 TACGACGTCGTTAAAGAGCTTACCTGAAATATTTTTCAGATTCGCGCGACGGCGTTAA 720  
 |||||  
 QY 994 tcacggagcgtcgagattcagcttgatcagcaacgagatgacgatttcgagattaca 1053  
 |||||  
 Db 721 TCACCGAGCTGGAGATCCCGCTCGAGACTCCGACGATGGTTTCAGATTCGCTATAACG 780  
 |||||  
 QY 1054 g 1054  
 |||||  
 Db 781 G 781

RESULT 4  
 LOCUS AW933524 643 bp mRNA linear EST 18-MAY-2001  
 DEFINITION EST359283 tomato fruit mature green, TAMU Lycopersicon esculentum CDNA clone cLEP54E2 5', mRNA sequence.  
 ACCESSION AW933524  
 VERSION AW933524.2 GI:11388127  
 KEYWORDS EST.  
 SOURCE tomato.  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
 Lycopersicon.  
 1 (bases 1 to 643)  
 Alcala, J., Vrebalov, J., White, R., Matern, A.L., Holt, I.E., Liang, F., Upson, J., Hansen, T., Craven, M.B., Bowman, C.L., Ahn, S., Ronaing, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.  
 Generation of ESTs from tomato fruit tissue

```
JOURNAL
COMMENT      Unpublished (1999)
              On May 30, 2000 this sequence version replaced gi:8108841.
              Contact: CUGI
              Clemson University Genomics Institute
              100 Jordan Hall, Clemson, SC 29634, USA
              Email: http://www.genome.clemson.edu/orders/index.html
              5 prime sequence.

FEATURES
  source
    1..643
      /organism="Lycopersicon esculentum"
      /cultivar="TA496"
      /db_xref="taxon:4081"
      /clone="cLEF54E2"
      /clone_lib="tomato fruit mature green, TAMU"
      /tissue_type="fruit pericarp"
      /dev_stage="mature green (3-5 days pre-ripening)"
      /lab_host="SOLR"
      /note="vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
      XhoI; cLEF - Fruit were tagged at the 1cm stage and
      harvested 3-5 days prior to ripening. Fruit were cut in
      half to verify the seeds were indeed 'immature' and the
      seeds and locules were discarded prior to freezing the
      pericarp"
    184 a 130 c 142 g 186 t 1 others
  BASE COUNT      184 a 130 c 142 g 186 t 1 others
  ORIGIN

Query Match      21.3%; Score 382.8; DB 9; Length 643;
Best Local Similarity 74.7%; Pred. No. 3.8e-103;
Matches 480; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

Qy 685 ttctacgcagagctgcagcgggtatagtcgaccggcacagcggcgtgtagctaac 744
Db 1  TTTACGCTCGTGGGCTCTTCGGACTTTGTATCACAGTAAGGAACCTGGTGTGCAAC 60

Qy 745 gccggtttggctctattcaatggcgggtattggctatgtcggaggatgattaccttac 804
Db 61 GCCGGTTTAGTCTATTTCAATAACCGATTACTTGTGTATGTCTGAAGATGATTGCCATTAC 120

Qy 805 caagttcagatcactcccaatggagatttaaaacccgttgctggttcgatttgatgga 864
Db 121 CATGTAAAGGTAACACCCACCGCGCATCTTAAACAGAGGGTCGATTGATTCGACGGC 180

Qy 865 caattagatccacaatgatgcaccaccggaagtgcgaccgggaatccggggaactcttc 924
Db 181 CAGCTAAAATCCACCATGATAGCTCACCCAAAGCTCGACCCAGTTTCGGGTGAGCTATT 240

Qy 925 gctttaagctacgagctggttcaagcgtctacctaataataactctcgattctcacccgac 984
Db 241 GCTCTTAGCTACGATGTGATTCAGAAGGCATACCTCAAGTACTTCAGATTTTCAAAAAT 300

Qy 985 ggaactaaatcacccgagctcgagatttcagcttgatcgagcgaacgatgcacgatttc 1044
Db 301 GGGGAAAATCAATGATGTTGAAATTCAGTTGAAGACCCCAACCAATGATGATGATTTC 360

Qy 1045 gcgattacagagaacttcgctcgtcgtacactgaccagcgaagtcggtttcaagctgcggag 1104
Db 361 GCAATTACTGAGAACTTCGTCGTCTATTCCTGATCAACAAGTCGTTTTCAAGATGCTGAA 420

Qy 1105 atgattcccggtgggtctccggtggtttacgacaagaacaagtcgcaagattcgggatt 1164
Db 421 ATGATCCGTGGAGGTTACCCGGTGTGTTTACGACAAGAACAAAGTTTCCGATTTGGTATT 480

Qy 1165 ttagacaataccgcgaagattcatcgaaacttaagtggtgattgatgtccgattgcttc 1224
Db 481 CTGGATAAGTACCGCAAGATGGGTCTGATTTGAAATGGGTTGAAGTACCTGATTGTTTC 540

Qy 1225 tgcctccatctctggaacgcttggaagagccagaaaacagatgaagtcgctcgtgatagg 1284
Db 541 TGTTTTCCACCTCTGGAATGCTTTGGGAAGAAGCAGAAAACAGATGAAATCGTTGTAATGGT 600

Qy 1285 tccgtgatgactcccacgactcaatttttcaacgagctcgacg 1327

Db 601 TCATGTATGACACCACCAGACTCCATTTTCAATGAATGTGATG 643

RESULT 5
BM408565
LOCUS
DEFINITION      BM408565 tomato breaker fruit Lycopersicon esculentum cDNA clone
                  EST 22-JAN-2002
                  787 bp mRNA linear
                  EST 22-JAN-2002
                  5' end, mRNA sequence.
                  cLEG45E23
                  5' end, mRNA sequence.
                  BM408565
                  5' end, mRNA sequence.
                  BM408565
                  5' end, mRNA sequence.
                  EST.
                  BM408565.1
                  GI:18260195
                  5' end, mRNA sequence.
                  tomato.
  ORGANISM
    Lycopersicon esculentum
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
    Lycopersicon.
  REFERENCE
    1 (bases 1 to 787)
    Alcala,J., Vrebalov,J., White,R., Vision,T., KaramyCheva,S.A., Tsai
    ,J., Bougri,O., Kirkness,E., Utterback,T., Van Aken,S., Roming
    ,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
    Generation of ESTs from tomato fruit tissue, breaker stage (2002)
    Unpublished (2002)
  CONTACT: CUGI
  CLEMSON UNIVERSITY GENOMICS INSTITUTE
  100 JORDAN HALL, CLEMSON, SC 29634, USA
  EMAIL: http://www.genome.clemson.edu/orders/index.html
  THIS CLONE IS AVAILABLE THROUGH THE CLEMSON UNIVERSITY GENOMICS
  INSTITUTE
  SEQ PRIMER: T3.
  LOCATION/QUALIFIERS
    1..787
      /organism="Lycopersicon esculentum"
      /cultivar="TA496"
      /db_xref="taxon:4081"
      /clone="cLEG45E23"
      /clone_lib="tomato breaker fruit"
      /tissue_type="pericarp"
      /dev_stage="breaker"
      /lab_host="SOLR"
      /note="Vector: pBluescriptSKmCuadapt; Site 1: EcoRI;
      Site 2: XhoI; supplier: Boyce Thompson Institute;
      sequencing: The Institute for Genomic Research. Fruit
      were harvested at the breaker stage (first sign of
      lycopenic accumulation on the blossom end of fruit). Fruit
      were cut in half and the seeds and locules were discarded
      prior to freezing the pericarp."
    216 a 194 c 165 g 212 t
  BASE COUNT      216 a 194 c 165 g 212 t
  ORIGIN

Query Match      20.9%; Score 375.6; DB 10; Length 787;
Best Local Similarity 67.8%; Pred. No. 6.2e-101;
Matches 525; Conservative 0; Mismatches 249; Indels 0; Gaps 0;

Qy 205 tctcccgcattgttgttaagcccaagcgaagaatccaaactaaacagatgaatttg 264
Db 5  TCACACCCAAAACAAAGAAAACACAACTCTCTTCTTCTTCAACTTCCAGTGGAAATTA 64

Qy 265 ttcagagagcggcgccgagcgttggacggcgagggtttccctgtcagccagag 324
Db 65 GTGCAGAAAGCAGCAGCAATGGCTTTAGTGTCTAGAAAGTGTCTTAACCTAAACATGAA 124

Qy 325 aagctacccgctctctaaacggctgatcctagtgctcagatcgccggaatttgc 384
Db 125 CTTGAACACCCCTTTGCCGAAAACAGCCGACCCAGAGTCCAGATTTCTGGGAAATTTGCT 184

Qy 385 ccggtgaatgaacggcgtccggcgtaattctccggtggtcggaactcccgattcc 444
Db 185 CCGGTACCGGAAAATCCAGTCTGTCAATCTCTTCGGTTCACCGGAAAATACCCAAATGT 244
```

QY 445 atcaaaaggagtgtatgtgcgaacggagctaaacccacttcacagacgggtgacaggteac 504  
 Db 245 GTTCAAGGGGTTTACGTTTCAAAACGGAGTAACCCCTCTTTTGAACCAACCCCGACAC 304  
 QY 505 cacttttcgacggagacggatggttcacgcgctcaaatcgaacacgggttcagctacg 564  
 Db 305 CATTCTTCGACGGGACGGTATGTTTCACGCGCTCAATTCATAAATGGTTCGGCTAGT 364  
 QY 565 tacgtttgcggtttactcagactaacacgggtttgttcaggaacgctcaattgggtcgacg 624  
 Db 365 TAGCGTTGCGGTTTCACTGAAACAGAGAGGCTTGTTCAGAAAGAAAGCTTTGGTTCGCCCT 424  
 QY 625 gttttcccaaaagcattcggttagcttcacgcccacacgggtatgcccagactcatgcta 684  
 Db 425 GTTTTCCCTTAAAGCCATTGGTGAATACATGCTACTCTGGAATTCGAAGGCTTATGCTG 484  
 QY 685 ttctacgcagagcgcagccggtatagtcgaccgggcacacggaacgggttagactaac 744  
 Db 485 TTTTACGCTCGTGGGCTCTTCGACTTGTGTGATCACAGTAAGGAAGACTGGTGTTCGAAC 544  
 QY 745 gccggtttgggtctatttcaatggcgggttattggctatgctcggaggtgatttacccttac 804  
 Db 545 GCCGGTTTAGTCTATTTCAATACCGATTACTTGTCTGTGAGATGATTTGCTTTAC 604  
 QY 805 caagttcagactcactcccaatggagatttataaacgggttggctcggttcgattttgatgga 864  
 Db 605 CATGTAAAGGTAAACACCCCGCGATCTTAAACAGAGGGTCGATTTCGATTTTCACGCGC 664  
 QY 865 caattagaatcaccaatgattccacccgaaagtcgaccgggaatccggtgaactcttc 924  
 Db 665 CAGCTAAATTCACCNATAGTCTACCCAAAGCTCGACCCGATTTCCCGGTGAGCTATT 724  
 QY 925 gctttaagctacgagctgtttcaagccttaccctaaatacttcgattctca 978  
 Db 725 GCTCTTAGCTACGATGTGATTCAGAAGCCATACCTCCAGTACTTCAGATTTTCA 778

RESULT 6  
 LOCUS BF113346 592 bp mRNA linear EST 18-MAY-2001  
 DEFINITION EST440936 tomato breaker fruit Lycopersicon esculentum cDNA clone  
 CLEG44B9 5' sequence, mRNA sequence.  
 ACCESSION BF113346  
 VERSION BF113346.1 GI:10943036  
 KEYWORDS EST.  
 SOURCE tomato.  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
 Lycopersicon.  
 REFERENCE 1 (bases 1 to 592)  
 AUTHORS Alcalá, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E.,  
 Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Ronning, C.M.,  
 Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley,  
 S.D.  
 TITLE Generation of ESTs from tomato fruit tissue, breaker stage  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>.  
 Location/Qualifiers  
 FEATURES  
 source  
 1..592  
 /organism="Lycopersicon esculentum"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CLEG44B9"  
 /clone\_lib="tomato breaker fruit"  
 /tissue\_type="pericarp"  
 /dev\_stage="breaker"  
 /lab\_host="SOLR"

/note="Vector: pBluescriptSKmUadapt; Site 1: EcoRI;  
 Site 2: XhoI; supplier: Boyce Thompson Institute;  
 sequencing: The Institute for Genomic Research. Fruit  
 were harvested at the breaker stage (first sign of  
 lycopene accumulation on the blossom end of fruit). Fruit  
 were cut in half and the seeds and locules were discarded  
 prior to freezing the pericarp."

BASE COUNT 194 a 106 c 131 g 161 t

Query Match 19.1%; Score 344; DB 10; Length 592;  
 Best Local Similarity 73.8%; Pred. No. 1.5e-91;  
 Matches 437; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 893 cgaagtcgacccggaatcccggtgaactcttcgctttaaagctacgacgtcggttcaaacg 952

Db 1 CAAAGCTCGACCCAGTTTCCGGTGAGCTATTGCTCTTAGCTACGATGTGATTGAGAAGC 60

QY 953 cttactaaaaatacttccgattctccacggagcgaactaaatcacccgacgtcgagattc 1012

Db 61 CATACCTCAAGTACTTTCAGATTTTCAAAAAATGGGAAAAATCAAAATGATGTTGAAATTC 120

QY 1013 agcttgatcagccaacgatgatgcagatttcgagatttcagagaaacttcgctcgctac 1072

Db 121 CAGTTGAAGACCCAAATGATGATGATTTTCGCAATTTACTGAGAACTTCGTCGTCATTTC 180

QY 1073 ctgaccagcaagtcgtttcaagctgccgagatgacccggtgggtctccggtgggttt 1132

Db 181 CTGATCAACAAGTCGTTTTCAAGATGCTGAAATGATCCGTGGAGGTTACCCGGTGGTTT 240

QY 1133 acgacaagaacaagtcgcaagattcgggattttagacaaataacccgaagattcatcga 1192

Db 241 ACGACAAGAACAAAGTTTCCCGATTGTTGTTTCTGGATAAGTACCGGAAAGATGGTCTCG 300

QY 1193 acattaaaggattgattcctcagatttctcctcctcctcctcctcctcctcctcctcctc 1252

Db 301 ATTTGAAATGGGTTGAAGTACCTGATTGTTTCTGTTTCCACCTCTGGAATGCTTGGGAG 360

QY 1253 agcagaacaagatgaagtcgctgctgataaggctcctgatactccaccagactcaattt 1312

Db 361 AAGCAGAAACAGATGAATCGTTGTAATGGTTCTATGATGATGATGATGATGATGATGAT 420

QY 1313 tcaacgagctctgacgagaatctcaagagtcctgtcgtgaaatccgctgaaatcctcaaaa 1372

Db 421 TCAATGAATGTGATGAAGGGCTAAAGAGTGTGTTTATCCGAAATCGTCTCAATTTGAAAA 480

QY 1373 ccggtgaatcaactccgctccgatcatctcccaagaaagatcaacaagtcacacctcgag 1432

Db 481 CAGGGAATCAACAAAGAAATCCATAATCGAAACCCCGATGAACCAAGTGAATTTAGAAG 540

QY 1433 caggatggtcaacagaacaatgctcggccgttaaaccaaaactcgttactt 1484

Db 541 CTGGAATGGTGAACCGAAACAAACTCGGAAGAAACAGAGTATGCTTATTT 592

RESULT 7

LOCUS BE461924

DEFINITION BE461924 566 bp mRNA linear EST 18-MAY-2001  
 clone CLEG40L2, mRNA sequence.

ACCESSION BE461924

VERSION BE461924.1 GI:9506322

KEYWORDS EST.

SOURCE tomato.

ORGANISM Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
 Lycopersicon.

REFERENCE 1 (bases 1 to 566)

Ahluwalia, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E.,  
 Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Ronning, C.M.,  
 Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley,  
 S.D.

Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley S.D.  
 Generation of ESTs from tomato fruit tissue, breaker stage  
 Unpublished (2000)  
 Contact: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

## FEATURES

Location/Qualifiers  
 source

1..566  
 /organism="Lycopersicon esculentum"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cUEG40L2"  
 /clone.lib="tomato breaker fruit, TIGR"  
 /tissue\_type="pericarp"  
 /dev\_stage="breaker"  
 /lab\_host="SOLR"  
 /note="Vector: pBluescriptSKMvUadapt; Site 1: EcoRI;  
 Site 2: XhoI; Fruit were harvested at the breaker stage  
 (first sign of lycopene accumulation on the blossom end  
 of the fruit). Fruit were cut in half and the seeds and  
 locules were discarded prior to freezing the pericarp."

BASE COUNT 184 a 104 C 124 G 154 T  
 ORIGIN

Query Match 18.5%; Score 332.4; DB 10; Length 566;  
 Best Local Similarity 74.2%; Pred. No. 4.3e-88;  
 Matches 420; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

Qy 893 cgaagtgcacccgaatcccggtgaactcttcgtttaagttacgaagtcgtttcaagc 952  
 Db 1 CAAGCTCGACCCAGTTCCGGTGAGCTATTGCTCTTAGCTACGATGTGATTCAGAAGC 60

Qy 953 cttaactaaatacttcgattctcaccggacgggaactaaatcaccggacgtcgagattc 1012  
 Db 61 CATACCTCAAGTACTTCAGATTTTCAAAAAATGGGAAAAATCAATGATGTGAAATTC 120

Qy 1013 agcttgatcagccaaacgatgatgcacgatcttcgggatttcacagaaacttcgtcgtgac 1072  
 Db 121 CAGTTGAGAGCCCAACATGATGATGATTTTCGCAATTTACTGAGNACTTCGTCGTCTTC 180

Qy 1073 ctgaccagagtcgttttcaagtcgccggagatgatccgcggtgggtctccggtggttt 1132  
 Db 181 CTGATCAACAAGTCGTTTTCAGATGTCAGAAATGATCCGTCGAGGTTTCACCGGTGGTGT 240

Qy 1133 acgacaagaacaagtcgaagattcgggattttagacaaatagccgaagattcatcga 1192  
 Db 241 ACGACAAGAACAAAGTTTCCCGATTGTGTAATCTGGATAAGTACCGGAAGATGGTCTG 300

Qy 1193 acattaaatgattgatctccagattctctcgtcttcattctctggaacgttgggaag 1252  
 Db 301 ATTGAAATGGTTGAAGTACCTGATGTTCTGTTTCCACTCTGGAATCTGTTGGGAG 360

Qy 1253 agcagaacagatgaagtcgtgtagaggttcctgtagtactccaccagactcaattt 1312  
 Db 361 AAGCAGAAACAGATGAATCGTTGTAATGGTTTCATGTATGACACACACAGACTCCATT 420

Qy 1313 tcaacgactcgacgagaaatctcaagatgctcctgctgaaatccgctgaatctcaaaa 1372  
 Db 421 TCAATGAATGTGATGAAGGGCTAAAGATGTTTATTCGAAATCCGTCATCAATTTGAAAA 480

Qy 1373 ccggtgaatcaactcgccgtccgcatctcctcaacgaagatcaacaagtcacacctcgaag 1432  
 Db 481 CAGGGAATCAACAGAAAAATCCATAATCGAAAAACCCGGATGACAAAGTATTTAGAG 540

Qy 1433 cagggatggtcaacagaacatgcttc 1458  
 Db 541 CTGGAATGTTGAACCGAAACAAACTC 566

## RESULT 8

BM085672

LOCUS

DEFINITION

BM085672

553 bp mRNA linear EST 19-NOV-2001

saJ28a02.y1 Gm-cl066 Glycine max cDNA clone SOYBEAN CLONE ID:

Gm-cl066-3867 5' similar to TR:O24023 O24023 NEOXANTHIN CLEAVAGE

ENZYME: ; mRNA sequence.

BM085672

GI:16996300

EST.

soybean.

Glycine max

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 553)

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna

, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,

Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers

, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk

, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann

, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Unpublished (1999)

Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntville, AL 35801 For further information

call: (800)-533-4363 or contact: [ccu@resgen.com](mailto:ccu@resgen.com) web site:

[www.resgen.com](http://www.resgen.com)

High quality sequence stop: 426.

Location/Qualifiers

1..553

/organism="Glycine max"

/db\_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl066-3867"

/clone.lib="Gm-cl066"

/tissue\_type="Leaf and shoot tip, salt stressed, 2 week

old seedling"

/lab\_host="DH10B"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI; The cDNA library was constructed from mRNA isolated

from unexpanded leaves and the shoot tips of 2 week old

seedling from the cultivar Williams. The 2 week old

seedlings were salt stressed in a solution of 500mM NaCl

for 3 days prior to harvesting. Complementary DNA was

synthesized from mRNA using a primer consisting of a

poly(dT) sequence with a XhoI restriction site. EcoRI

adapters were ligated to the blunt-ended cDNA fragments

followed by XhoI digestion. The cDNA fragments were

directionally cloned into the EcoRI-XhoI restriction site

of the pBluescript vector. The ligated cDNA fragments were

transformed into DH10B host cells (GibcoBRL). This library

was constructed in the laboratory of Dr. Randy

Shoemaker."

BASE COUNT 148 a 165 C 120 G 120 T

ORIGIN

Query Match 18.0%; Score 323.2; DB 10; Length 553;

Best Local Similarity 74.1%; Pred. No. 2.4e-85;

Matches 409; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

Qy 822 caatggagatttaaaacccgttggtcggttcgattttgatggacaattagaatcaaat 881

Db 1 CAACGGGAGCTTAACACCGCTCGCGCTTACAACTTTAACGGCCAGTTAAATCACAAT 60



RESULT 10  
 BE432853  
 LOCUS  
 DEFINITION EST39478 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA  
 clone cLEG11G7, mRNA sequence.  
 ACCESSION BE432853  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
 Lycopersicon.  
 1 (bases 1 to 520)  
 Alcala, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E.,  
 Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Ronning, C.M.,  
 Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley  
 , S.D.  
 TITLE Generation of ESTs from tomato fruit tissue, breaker stage  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..520  
 /organism="Lycopersicon esculentum"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cLEG11G7"  
 /clone\_lib="tomato breaker fruit, TIGR"  
 /tissue\_type="pericarp"  
 /dev\_stage="breaker"  
 /lab\_host="SOLR"  
 /note="Vector: pBlueScriptSKmCudapt; Site 1: EcoRI;  
 Site 2: XhoI; Fruit were harvested at the breaker stage  
 (first sign of lycopene accumulation on the blossom end  
 of the fruit). Fruit were cut in half and the seeds and  
 locules were discarded prior to freezing the pericarp."  
 BASE COUNT 156 a 100 c 116 g 148 t  
 ORIGIN  
 Query Match 17.6%; Score 316.6; DB 10; Length 520;  
 Best Local Similarity 75.9%; Pred. No. 2.2e-83;  
 Matches 391; Conservative 0; Mismatches 124; Indels 0; Gaps 0;  
 Qy 829 gatttaaaacccgttgctggttcatttggatggacaaattagatcacacaattgattgcc 888  
 Db 2 GATCTTAAACAGAGGGTGGATTTCGATTCGACGGCCAGCTAAATCCACCATGATAGCT 61  
 Qy 889 caccgaaagtcgaccggaatccggtgaactcttcgctttaagctacgagctgttca 948  
 Db 62 CACCCAAAGCTCGACCCAGTTTCCGGTGAGCTATTTCCTTTAGTACGATGTAATCAG 121  
 Qy 949 aagccttaactaaaatactccgattctcaccggcgagcgaactaaatcacggagctcgag 1008  
 Db 122 AAGCCATACCTCAAGTACTTCAGATTTTCAAAAATGGGAAAAATCAATGATGTTGAA 181  
 Qy 1009 attcagcttgatcagccaacgatgatgacgatttcgcgattacagagaacttcctgc 1068  
 Db 182 ATTCCAGTTGAAGACCCCAACAATGATGATGATGATTTTCGCAATTACTTGAGAACTTCGTCGTC 241  
 Qy 1069 gtacctgaccagcaagtcgttttcgaagtcgcgagatgatccgggtgggtctccggtg 1128  
 Db 242 ATTCTGTATCAACAAGTCGTTTTCAAGATGCTGAAATGATCCGGTGGAGGTTTACCCGGTG 301  
 Qy 1129 gtttacgacaagaacaaagtcgcaagattcgggatttttagacaaataacgccgaagtcca 1188  
 Db 302 GTTTACGACAAGAACAAAGTTTCCCGATTGTTGTTATTTCTGGATAAGTACGCGAAAGATGGG 361

Qy 1189 tcgaacattaaagtgattgatgtccagattgcttctgcttccatctctctgaacgcttg 1248  
 Db 362 TCTGATTGAAATGGTTGAAGTACCTGATGTTTCTGTTTCCACCTCTGGAATGCTTGG 421  
 Qy 1249 gaaagccagaacagatgaagtcgtcgatagggctcctgtatgactccaccagactca 1308  
 Db 422 GAAGAAGCAGAAACAGATGAATCGTTGTAATTGGTTTCATGTATGACACCAACCACTCC 481  
 Qy 1309 attttcaacagtgctgacgagaatctcaagagtg 1343  
 Db 482 ATTTTCAATGAATGTGATGAAGGCTAAAGAGTGT 516  
 RESULT 11  
 BH559796  
 LOCUS  
 DEFINITION BOHR47TF BOHR Brassica oleracea genomic clone BOHR47, DNA  
 sequence.  
 ACCESSION BH559796  
 VERSION BH559796.1 GI:17811576  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea.  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 742)  
 Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.  
 TITLE Whole genome shotgun sequencing of Brassica oleracea  
 JOURNAL Unpublished (2001)  
 COMMENT Other GSSs: BOHR47TR  
 Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: [cdtown@igr.org](mailto:cdtown@igr.org)  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TP  
 Class: Sheared ends.  
 FEATURES  
 Location/Qualifiers  
 source  
 1..742  
 /organism="Brassica oleracea"  
 /strain="TOL000DH3"  
 /db\_xref="taxon:3712"  
 /clone="BOHR47"  
 /clone\_lib="BOHR"  
 /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared  
 genomic DNA inserted into pHOS1 using BstXI linkers"  
 BASE COUNT 190 a 235 c 117 g 200 t  
 ORIGIN  
 Query Match 17.4%; Score 313.8; DB 12; Length 742;  
 Best Local Similarity 84.1%; Pred. No. 1.9e-82;  
 Matches 354; Conservative 0; Mismatches 67; Indels 0; Gaps 0;  
 Qy 1 atggctttttcacggcaacggctgcgggtttctcggagatggcttggtggcaatcatact 60  
 Db 371 ATGGCTTCTTTTCACGGGACTACGGCGGTTTCTCGAGATGGTGGTAAATCATACT 380  
 Qy 61 cagcgcgcattatcgctttctcaagctccgacttgagttattgtaccttacctatg 120  
 Db 381 AAAGCACCATTTATCGTCTTCTTCAAAGTCCGCCCTTGGGTTATTCTTCGCTGCCATGCCATG 440  
 Qy 121 gccagtcgtgcacacgaagtcgaatttcattctcgcttcacactcctccagctctt 180  
 Db 441 ACAATCGTTCCTCCCAACGAAAGCTCAATGTTTCTTCTCGGCTTCACACTCATCTCTGCTCTC 500  
 Qy 181 catttccctaagcaatcatcaaatctcccgccatttctgtaagcccaagccaaagaa 240  
 Db 501 CATTTCCCAAGCAATCTCCACCTCTCCCGCCATTGTTGTGAACCCCAACCAAGAA 560



greenhouse grown"  
/lab host="DH10B"  
/note=vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from drought stressed leaf tissue of the cultivar Williams  
82. The month old greenhouse grown plants were deprived of  
water for 3 days prior to harvesting the stressed leaf  
tissue. Complementary DNA was synthesized from mRNA using  
a primer consisting of a poly(dT) sequence with a XhoI  
restriction site. EcoRI adapters were ligated to the  
blunt-ended cDNA fragments followed by XhoI digestion. The  
cDNA fragments were directionally cloned into the  
EcoRI-XhoI restriction site of the pBluescript vector. The  
ligated cDNA fragments were transformed into DH10B host  
cells (GibcoBRL). This library was constructed in the  
laboratory of Dr. Randy Shoemaker."

```
BASE COUNT 123 a 141 c 133 g 150 t
ORIGIN
Query Match 17.2%; Score 309.4; DB 10; Length 547;
Best Local Similarity 74.4%; Pred. No. 3.2e-81;
Matches 404; Conservative 0; Mismatches 136; Indels 3; Gaps 1;

Qy 991 aaatcacggagctcgagattcagcttgatcagccaacgatgcacgatttcgagatt 1050
Db 541 AAATCCCCGAGCTCGAAATTCCTTGAAGAACCCCATGATGACGATTTGCCATC 482

Qy 1051 acagagaacttcgtctgcgtacccgaccagcaagtcggttttcaagctccggagatgac 1110
Db 481 ACCGAGAACTTCGTCGTGCCCGACCGAGGTGGTCTTCAAGTTATCAGAAATGATC 422

Qy 1111 cgcggtgggttcgggttttaagacaagaacaaagtcgcaagattcgggatttagac 1170
Db 421 ACCGAGGCTCCCCGTTGTCTACGACAAAGCAACAAAGGTTTCCAGATTCGGATTCTCGAC 362

Qy 1171 aaatcacggaagattcatcgaacatttaagtgattgattcgcagattgcttctgttc 1230
Db 361 AAGATGCTTAAAGATGCAAAATGATATGAATGGATCGATGCTCCGAGTGTCTTGTTTC 302

Qy 1231 catctctggaacgcttgggaagagcagaaacagatgaagtcgctgtagagggctcgt 1290
Db 301 CACCTCTGGAACGCTGGGAGAGCGGAAATGATGATGATGCTGTTCATCGGTTCTCTGC 242

Qy 1291 atgaattcacacagatcaatttttaacagagtcgacgagaatctcaagagtgctctgtct 1350
Db 241 ATGACCCCCGGGATCCCAATTTCAACGAATCCGAAGAGAGTTTGAAGAGCATCTTGTCG 182

Qy 1351 gaaatccgcctgaattctcaaacccggtgaatcaactcgcgcgcgatcatctccaacgaa 1410
Db 181 GAGATAAGGCTGAATTTGAAGACAGGCAAGTCCACGAGAAACCCATTTATCTC---GGAA 125

Qy 1411 gatcaacaagctcaactcgaagcaggatggttcaacagaaacatgctcggcgtaaaacc 1470
Db 124 TCCGAAACAAGTCAACTTGGAAAGCCGGATGGTCAACAGAAACAAGCTCGGGAGAAAGAGC 65

Qy 1471 aaattcgcttacttggttttagccgagccgtggcctaaagctcaggattcgctaaagtt 1530
Db 64 AAGTTTCGGGTACTTAGCACATTGCGGAGCCGTGTCTTAAGGTTTCGGGTTTTCGCCAAGGTT 5

Qy 1531 gat 1533
Db 4 GAT 2
```

RESULT 14  
BE437072  
LOCUS BE437072 547 bp mRNA linear EST 18-MAY-2001  
DEFINITION EST408190 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA  
clone cLEG35K8, mRNA sequence.  
ACCESSION BE437072  
VERSION BE437072.1 GI:9434915  
KEYWORDS EST.

SOURCE  
ORGANISM

tomato.  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.  
1 (bases 1 to 547)  
Alcala J., Vrebalov J., White, R., van der Hoeven, R.S., Holt, I.E.,  
Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Roming, C.M.,  
Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley  
S.D.

Generation of ESTs from tomato fruit tissue, breaker stage  
Unpublished (2000)

CONTACT: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES

Location/Qualifiers

source

1..547  
/organism="Lycopersicon esculentum"  
/cultivar="TA96"  
/db xref="taxon:4081"  
/clone="cLEG35K8"  
/clone\_lib="tomato breaker fruit, TIGR"  
/tissue\_type="Pericarp"  
/dev\_stage="breaker"  
/lab\_host="SOLR"  
/notes="Vector: pBluescriptSKmCuadapt; Site 1: EcoRI;  
Site 2: XhoI; Fruit were harvested at the breaker stage  
(first sign of lycopene accumulation on the blossom end of  
the fruit). Fruit were cut in half and the seeds and  
locules were discarded prior to freezing the pericarp."

BASE COUNT 140 a 133 c 120 g 154 t  
ORIGIN

Query Match

17.1%; Score 308.6; DB 10; Length 547;

Best Local Similarity 72.8%; Pred. No. 5.6e-81;

Matches 398; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

Qy 431 aactcccgattccatcaaaaggagtgatgtgcgaacgagctaaacccacttcacgagc 490

Db 1 AAATACCAAAATGTTCAAGCGGTTTACGTTCCAAACGGAGCTAACCTCTTTTGAAC 60

Qy 491 cggtgacaggtcaccaactcttcgacggagacggtatgggttcacgcggtcaaatcgaac 550

Db 61 CAACCGCGGACACCAATTTCTTCACGCGGACGGTATGGTTACGCGGTTCAATTCAAAA 120

Qy 551 acggttcagctagctagcttgcgggtttactcagactaacgggttgggttcaggaaagtc 610

Db 121 ATGGGTGGCTAGTTTACGCTTTCCTGAAACAGAGAGGCTTGTTCAGAAAAAG 180

Qy 611 aattgggtcgaccggttttccccaaagccatcggtgagcttcacggccacacccggtattg 670

Db 181 CTTTGGGTGGCCCTGTTTCCCTAAAGCCATTGGTGAATACATGGTCACTCTGGAATTG 240

Qy 671 cccgactcgtctattctacgcccagagctgcagccggtatagtcgaccccgccacacgaa 730

Db 241 CAAGGCTTATGCTGTTTACGCTCGTGGCTCTTCGGACTTGTTCATCACAGTAAGAA 300

Qy 731 ccggtgtagctaacgcccgggttggctatttcaattgcccgggttattggctatgctggagg 790

Db 301 CTGGTGTGCAACCGCGGTTTAGTCTATTTCAATAACCGGATTACTTGTCTATGCTGAAG 360

Qy 791 atgatttaccttaccaggttcagatcactcccaatgagatttaaaaccccttggctcgtt 850

Db 361 ATGATTTGCCCTTACATGTAAGGTAAACACCCCGCGGATCTTTAAACAGAGGTCGAT 420

Qy 851 tcgattttgatggacaattagaatcccaatgattgcccaccccgaaagtcgaccgggaat 910

Db 421 TCGATTTCCGACGGCCAGCTAAAAATCCACCATGATAGCTCACCCAAAGCTCGACCCAGTTT 480



```

QY 911 cgggtgaactcttcggtttaagctacagcgtcgcttcaaaagccttaccctaaataacttc 970
Db 481 CCcGTGAGCTATTGTCTTAGCTACGATGTGATTcAGAAGCCATACCTCAAGTACTTCA 540
QY 971 gattctc 977
Db 541 GATTCTC 547

RESULT 15
BE458861 BE458861 617 bp mRNA linear EST 18-MAY-2001
LOCUS EST14153 tomato developing/immature green fruit Lycopersicon
DEFINITION esculentum cDNA clone cLEM4F10, mRNA sequence.
ACCESSION BE458861
VERSION BE458861
KEYWORDS EST.
SOURCE BE458861.1 GI:95031163
ORGANISM tomato.
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 617)
AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
,S.D.
TITLE Generation of ESTs from tomato fruit tissue, immature green
JOURNAL Unpublished (2000)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES
    Location/Qualifiers
        1..617
            /organism="Lycopersicon esculentum"
            /cultivar="TA496"
            /db_xref="taxon:4081"
            /clone="cLEM4F10"
            /clone_lib="tomato developing/immature green fruit"
            /tissue_type="fruit"
            /dev_stage="immature green (5-35 days post-anthesis)"
            /lab_host="SOLR"
            /note="Vector: pBluescriptSKmCuadapt; Site 1: EcoRI;
            Site 2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and
            harvested at 7 day intervals through 35 dpa. Equal masses
            of tissue from each stage were combined (including seeds
            and locules) prior to mRNA isolation."
BASE COUNT 164 a 145 c 137 g 171 t
ORIGIN

Query Match 16.9%; Score 303.4; DB 10; Length 617;
Best Local Similarity 68.6%; Pred. No. 2.2e-79;
Matches 418; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 247 actaacagatgaattgttcacagagcggcgggcggttgagcggcgagggt 306
Db 9 ACTTCCAAGTGGAAATTTAGTCAGAAAGCAGCAGCAATGGCTTAGATGCTGTAGAAAGT 68
QY 307 ttcctgtcagcagagaagatacccgcttcctaaacggctgactcgtgttcag 366
Db 69 GGTTTAACTAAACATGAACCTTGAACACCCCTTTGCCGAAAACAGCCGACGAGTCCAG 128
QY 367 atcgccggaaatttgcctccgtgaatgaacggccgctccggtaatcttcgggtggtc 426
Db 129 ATTTCTGGGAATTTGCTCCGGTACCGGAAATCCAGTCTGTCAATCTCTTCGGTCACC 188
QY 427 ggaaacttccgattccatcaaaaggagtgtatgtgcgaacggagtaacccacttcac 486

```

```

Db 189 GGAATAATACCCAAATGTGTTCAAGGCGTTTACGTTTCAAAACGGAGCTAACCCCTCTTTT 248
QY 487 gggcggtgacaggtcaccaactcttcgacgagacggtatggttcacgcggtcaaatc 546
Db 249 GAACCAACCCCGGACACCATTTCTTCGACGGGACGGTATGTTCCACCCGTTCAATTC 308
QY 547 gaacacggttcagctacgtacgcttgccgggtttactcagactaacccgggtttgttcaggaa 606
Db 309 AAAAATGGGTGGCTAGTTACGCTTGCCGTTTCACTGAAACAGAGAGGCTTTGTTCAAGAA 368
QY 607 cgtcaatgggtcgaccgggttttccccaaagccatcggtgagcttcacggccacacccgt 666
Db 369 AAAGCTTTGGGTGCGCCCTGTTTCCCTAAAGCCATTTGGTGAATTACATGCTCACTCTGA 428
QY 667 attgcccgaactcatgctattctacgcccagagctgcagccggtatagtcgacccggcacac 726
Db 429 ATTGCAAGGCTTATGCTGTTTACGCTCGTGGGCTCTTCGGACTTGTGATCACAGTAA 488
QY 727 ggaaccggtgtagctaacgcccgggtttggtctatttcaatggccggttattggctatgtcg 786
Db 489 GGAACCTGGTGTTCGAAACGCCGCTTTAGTCTATTTCATTAACCGATTACTTGTATGTCT 548
QY 787 gaggatgatttacccttaccgaagttcagatcactcccaatggagatttaaaacccgttgt 846
Db 549 GAAGATGATTTTCCTTACCATGTAAGGTAACACCCACCGCGCATCTTAAACAGAAAGT 608
QY 847 cggttccat 855
Db 609 CGATTCCGAT 617

```

Search completed: July 24, 2002, 06:53:47  
Job time: 5972 sec

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: November 15, 2003, 09:15:54 ; Search time 6156 Seconds  
(without alignments)  
3980.648 Million cell updates/sec

Title: US-09-758-269-6  
Perfect score: 3150  
Sequence: 1 MASFTATAAASGRWLGNHT.....VPYGFHGTFIGADDLAKQVV 599

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2889711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US09758269/runat.14112003.135756.28745/app.query.fasta.1.775  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09758269 @CEN 1.1 3963 @runat.14112003.135756.28745 -NCPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3150	100.0	1800	6	AX148306 Sequence
2	3150	100.0	1800	6	BD017431 Transgeni
3	3150	100.0	3869	8	AB026549 Arabidops
4	3150	100.0	52232	8	AB028617 Arabidops
5	3143	99.8	2331	8	AY056255 Arabidops
6	2291	72.7	2164	8	STU276244 Solanum t
7	2282.5	72.5	2171	8	LENEOXANT
8	2282.5	72.5	19018	8	LES439079 Lycopersi
9	2280.5	72.4	1818	6	AX148316 Sequence
10	2172	69.0	2398	8	AF190462 Phaseolus
11	2170.5	68.9	95769	6	AX652128 Sequence
12	2170.5	68.9	95769	6	AC013430 Genomic s
13	2168.5	68.8	1839	6	AX148312 Sequence
14	2168.5	68.8	1839	6	BD017434 Transgeni
15	2168.5	68.8	2432	8	AB030293 Vigna ung
16	2155.5	68.4	2357	8	AB080192 Pisum sat
17	2075.5	65.9	2034	8	AB080193 Pisum sat
18	2066.5	65.6	2310	8	AF224671 Persea am
19	2044.5	64.9	72058	8	AC074176 Arabidops
20	1991	63.2	1752	6	AX148302 Sequence
21	1991	63.2	1752	6	BD017429 Transgeni
22	1991	63.2	110102	8	ATF28J12
23	1991	63.2	110102	8	ATCHRIV48
24	1991	63.1	175158	2	AC120531 Oryza sat
25	1988	63.1	181159	2	CNS071P7
26	1987	63.1	2153	8	AF224672 Persea am
27	1959.5	62.2	133983	8	CNS08CB9
28	1959.5	62.2	131689	8	CNS08CB9
29	1930	61.3	1815	6	AX148314 Sequence
30	1930	61.3	1815	6	BD017435 Transgeni
31	1930	61.3	2498	8	ZMU095953
32	1834.5	58.2	149097	2	AP005632 Oryza sat
33	1663.5	52.8	1734	6	BD017433 Transgeni
34	1663.5	52.8	1734	6	AB028621 Arabidops
35	1663.5	52.8	78921	8	AB028621 Arabidops
36	1663.5	52.8	82697	8	AC123596 Medicago
37	1214	38.5	120055	8	AC123596 Medicago
38	1014.5	32.2	120055	8	AC123596 Medicago
39	1003	31.8	80686	2	AC144759 Medicago
40	986.5	31.3	120055	8	AC123596 Medicago
41	986.5	31.3	2016	8	AB080194 Pisum sat
42	971.5	30.8	80686	2	AC144759 Medicago
43	969	30.8	1788	6	AX148304 Sequence
44	969	30.8	1788	6	AX506749 Sequence
45	969	30.8	1788	6	BD017430 Transgeni

ALIGNMENTS

RESULT 1

```

AX148306
LOCUS AX148306 1800 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 5 from Patent EP1116794.
ACCESSION AX148306
VERSION AX148306.1 GI:14347193
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE Iuchi, S., Kobayashi, M. and Shinozaki, K.
AUTHORS Transgenic plants carrying neoxanthin cleavage enzyme gene
TITLE Patent: EP 1116794-A 5 18-JUL-2001;
JOURNAL Riken (JP)
FEATURES
source
location/Qualifiers
1..1800
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
1..1800
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC41194.1"
/db_xref="GI:14347194"
/translation="MASFTATAVSGRWLGNHTQPPLSQSSDLSYCSLPLMASRV
TRKUNVSAHLTPALHPKQSSNSPAIVVKPAKESNTKOMNLFORAAAALDAAGS
FLVSEKHLPLKTPADPSVQIAGNAPVNEQPVRRNLPVVGKLPDSIKGVYVRNGANP
LHEPVTGHHFPGDGMVHAKVFHFSGSAYACRFQTNRFVQEROLGRPVFPKAIQELH
GHTGIARLMLFYARAAGIIVDPAGHTGVNAGLVFNGRLLAMSEDDL PYQVQITPNG
DLKTVGRFDDQQLSEMIAPKVPDPSGELFALSVDYVSKPYLKYFRFPDGTKSPD
VEIQDQTMHDFATENTFVVDPQVVFKLPEMIRGSGSPVVDKMKVAFGLIDKY
AEDSSNTKIDAPDFCFHLMNAWEEPETDEVVVGISMTTPPDSIFNEDENLKSULS
EIRLNTKGTSESTRPIISNEDQVNEAGVMVNRMLGRKTFAYIALAEWPVKVSGFA
KVDLTTEVKKHLYGDMRYGGEPLFLPGEGEDEDEYILCFVHDEKTKWSELQIVNAV
SLEVEATVKLPSRYPVGFHGTFFIGADDLAKQVV"
BASE COUNT 458 a 464 c 439 g 439 t
ORIGIN
Alignment Scores:
Pred. No.: 6,73e-220 Length: 1800
Score: 3150.00 Matches: 599
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-758-269-6 (1-599) x AX148306 (1-1800)
Qy 1 MetAlaSerPheThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThr 20
Db 1 ATGGCTTCTTTTACGGCAACCGCTCGCGTTCCTGGAGATGGCTTGGTGGCAATCATACT 60
Qy 21 GlnProLeuSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet 40
Db 61 CAGCGCCCATATTCGTCTCTTCAAAGCTCCGACCTTGAGTTAGTTAGTCTCTTACCTATG 120
Qy 41 AlaSerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProAlaLeu 60
Db 121 GCCAGTCGTGTACACAGTAAGCTCAATGTTTTCATCTGCGCTTTCACACTCTCTCCAGCTCTT 180
Qy 61 HisPheProLysGlnSerSerAsnSerProAlaLeuValValLysProLysAlaLysGlu 80
Db 181 CATTTCCCTAAGCAATCATCAAACTCTCCGCCCATTTGTTGTTAAGCCCAAGCAAGAA 240
Qy 81 SerAsnThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaLeuAspAlaAla 100
Db 241 TCCAACTAACTAAACAGATGAATTTGTTCCAGAGACGGCGCGGAGCGCTTGGACGGCGC 300
Qy 101 GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer 120
Db 301 GAGGGTTTCTTGTGTCAGCCACGAGAGCTACACCGCTTCTCTAAACCGGCTGATCTCTAGT 360

```

---

```

121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuPro 140
361 GTTCAGATCGCGGAAATTTTGTCTCCGGTGAATGAACAGCCCGTCCGGCGTAATCTTCCG 420
141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160
421 GTGGTCGGAATACTTCCGATTCATCAAGAGGAGTGATGTGCGCAACGAGCTTAACCCA 480
161 LeuHisGluProValThrGlyHisHisPhePheAspGlyAspGlyMetValHisAlaVal 180
481 CTTTACGAGCGGTGACAGGTCACCACTTCTTCGACGAGACGATGATGGTTCACGCGTC 540
181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200
541 AAATTCGAACACCGTTTCAGCTAGCTACGCTTTCGCGGTTTACTCAGATAACCGGTTTGT 600
201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220
601 CAGGAACGTCAATTGGGTGACCGGTTTCCCAAGCCCATCGGTGAGCTTCACGGCCAC 660
221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspPro 240
661 ACCGTTATTCGCCACTCATGCTATTCTAGCCAGAGCTGCAGCGGTATAGTCGACCCG 720
241 AlaHisGlyThrGlyValAlaAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuAla 260
721 GCACACGGAACCGGTGAGCTAACCGCGTTGGTGTCTATTTCATATGCGCGGTATTGGCT 780
261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280
781 ATGTCGAGAGATGATTTACCTTACCAAGTTTCAGATCACTCCCAATGGAGATTTAAAAACC 840
281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300
841 GTTGGTCGGTTCGATTTGATGACCAATTAGAATCCACAATGATTGTCGCCACCCGAAAGTC 900
301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeu 320
901 GACCCGGAATCCCGTGAACTCTTCGCTTTAAGCTTACGACGTCGTTTCAAAGCCTTACCTA 960
321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340
961 AATATCTTCCGATCTTCACCGGACGGAATTAATCACCGGACGTCGAGATTCAGCTTGTAT 1020
341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValValProAspGln 360
1021 CAGCCAAACGATGATGCACGATTTTCGCGATTACAGAGAATTCGTCGTCGATGACCCAG 1080
361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380
1081 CAAGTCGTTTTCAGCTGCGCGAGATGATCCGCGTGGGTCTCCGGTGGTTTACGACAAG 1140
381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400
1141 AACAGGTCGCAAGATTCGGGATTTTAGACAATATACCGCCGAGATTCATCGAATTAAG 1200
401 TrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluProGlu 420
1201 TGGATTGATGTCAGATTCGCTTCTGCTTCATCTCTGGAACGCTTGGGAAGAGCCAGAA 1260
421 ThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440
1261 ACAGATGAAGTCGTGATGAGGTCTCTGTATGACTCCACAGACTCAATTTTCAACGAG 1320
441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460
1321 TCTGACAGAGATCTCAAGAGTGTCCTGTCTGAAATCCGCTGAATCTCAAAACCGGTGAA 1380
461 SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480
1381 TCAACTCGCGTCCGATCATCTCCACGAAAGATCAACAAGTCAACCTCGAAGCAGGATG 1440

```

QY 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500  
 DB 1441 GTCACAGAAACATCTCGCGCGTAAAAACCAATTCGCTTACTTGCGCTTTAGCCGAGCG 1500  
 QY 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis 520  
 DB 1501 TGGCCTTAAGTCTCAGGATTCGTAAGTTGATCTCCTACTCGGAGAAGTTAAGAAACAT 1560  
 QY 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGlyGlu 540  
 DB 1561 CTTTACGGCGATAACCGTTACGGAGAGAGCCCTCTGTTCTCCCGGAGAGGAGAG 1620  
 QY 541 GluAspGluGlyTyrLysLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560  
 DB 1621 GAAGACGAAGGATACATCTCTGTTTCGTTTCACGACGAGACATGGAATCGGAGTTA 1680  
 QY 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580  
 DB 1681 CAGATAGTTAAGCCGTTAGCTTAGAGTTGAGGCAACGGTTAACTTCCGTCAGGGTT 1740  
 QY 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGlnValVal 599  
 DB 1741 CCGTACGGATTTACGGTACATTCATCGAGCGGATGATTTGGCGAAGCAGGTCGTG 1797

## RESULT 2

BD017431 1800 bp DNA linear PAT 27-AUG-2002  
 LOCUS Transgenic plant using neozanthine cleaving enzyme gene.  
 DEFINITION  
 ACCESSION BD017431  
 VERSION BD017431.1 GI:22558607

KEYWORDS JP 2001258579-A/3.

SOURCE Arabidopsis thaliana (thale cress)

## ORGANISM

Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

## REFERENCE

1 Iuchi, K., Kobayashi, M. and Shinozaki, K.  
 Transgenic plant using neozanthine cleaving enzyme gene  
 Patent: JP 2001258579-A 3 25-SEP-2001;  
 THE INSTITUTE OF PHYSICAL AND CHEMICAL RESEARCH  
 OS Arabidopsis thaliana (thale cress)  
 PN JP 2001258579-A/3

PD 25-SEP-2001

PF 11-JAN-2001 JP 2001003476

PI KIYOSHI IUCHI, MASATOMO KOBAYASHI, KAZUO SHINOZAKI PC

C12N15/09,A01H5/00,C12N5/10,C12N5/02,C12N5/00,C12N5/00 CC

Transgenic plant using neozanthine cleaving enzyme gene PH Key

Location/Qualifiers (1)..(1800).

## FEATURES

source

1..1800

/organism="Arabidopsis thaliana"

/mol\_type="genomic DNA"

/db\_xref="taxon:3702"

BASE COUNT 458 a 464 c 439 g 439 t

## ORIGIN

Alignment Scores:

Pred. No.: 6,738-220 Length: 1800  
 Score: 3150.00 Matches: 599  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-758-269-6 (1-599) x BD017431 (1-1800)

QY 1 MetAlaSerPheThrAlaThrAlaValSerGlyArgTrpLeuGlyLysHisThr 20

DB 1 ATGGTCTTTCACGGCAACGGTCGGGTTCTCGGAGATGGCTTGGTGGCATCATACT 60

QY 21 GlnProProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet 40

DB 1 ATGGTCTTTCACGGCAACGGTCGGGTTCTCGGAGATGGCTTGGTGGCATCATACT 60

QY 21 GlnProProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet 40

DB 1 ATGGTCTTTCACGGCAACGGTCGGGTTCTCGGAGATGGCTTGGTGGCATCATACT 60

QY 21 GlnProProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet 40

DB 1 ATGGTCTTTCACGGCAACGGTCGGGTTCTCGGAGATGGCTTGGTGGCATCATACT 60

QY 21 GlnProProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet 40

DB 1 ATGGTCTTTCACGGCAACGGTCGGGTTCTCGGAGATGGCTTGGTGGCATCATACT 60

DB 61 CAGCGCGCATTTATCGTCTTCTCAAAGCTCCCAAGCTTGAAGTTATTTAGTCTCTTACCTATG 120  
 QY 41 AlaSerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProProAlaLeu 60  
 DB 121 GCCAGTCGTGTACACGTAAGCTCAATGTTTCATCTGGGCTTCACACTCTCTCCAGCTCTT 180  
 QY 61 HisPheProLysGlnSerSerAsnSerProAlaIleValValLysProLysAlaLysGlu 80  
 DB 181 CATTTCCCTAAGCAATCATCAAACTCTCCCGCCATTGTTGTTAAGCCCAAGCCAAAGAA 240  
 QY 81 SerAsnThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaAlaAlaAspAlaAla 100  
 DB 241 TCCAACACTAAACAGATGAATTTGTTCCAGAGAGCGCGCGGAGCGTTGGACCGCGG 300  
 QY 101 GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer 120  
 DB 301 GAGGGTTTCCTTGTCGCCACGAGAGCTACACCGCTTCTCTAAAACGGCTGATCCTAGT 360  
 QY 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgAsnLeuPro 140  
 DB 361 GTTCAGATCGCGGAAATTTTGCTCCGGTGAATGAACAGCCGTCGCGCGTAATCTTCGG 420  
 QY 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160  
 DB 421 GTGGTCGGAACACTTCCGATTCCATCAAGAGAGTGTATGTGCGCAACGGAGCTAACCCA 480  
 QY 161 LeuHisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaVal 180  
 DB 481 CTTACAGACCGGTGACAGGTCACCACTTCTTCAGCGGAGACGGTATGGTTACCGCCGTC 540  
 QY 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200  
 DB 541 AAATTCGAACACGGTTTACGTTAGCTAGCTAGCTTTCGCGGTTTACTCTACAGCTAACCGGTTGTT 600  
 QY 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLysHisGlyHis 220  
 DB 601 CAGGAACGTCAAATTCGGGTTCGACCGGTTTCCCAAGGCAATCGGTGAGCTTTCACGCGCAC 660  
 QY 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspPro 240  
 DB 661 ACCGGTATTCGCCACTCATCTATTCTACGCCAGAGCTGCAGCCGGTATAGTCACGCCG 720  
 QY 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260  
 DB 721 GCACACGAACCGGTGTAGCTAACCGCGTTTGGTCTATTTCAATGGCCGGTTATTGGCT 780  
 QY 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280  
 DB 781 ATGTCGGAGGATGATTTTACCTTACCAAGTTCAGATCACTCCCAATGGAGATTTAAAAACC 840  
 QY 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300  
 DB 841 GTTGGTCGTTTCGATTTTGTATGGCAATTAGAATCCACAAATGATTCGCCACCCGGAAGTC 900  
 QY 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeu 320  
 DB 901 GACCCGGAATCCGGTGAACCTCTCGCTTAAAGCTACGACGTCGTTTCAAGAGCTTACCTA 960  
 QY 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340  
 DB 961 AAATACCTTCCGATTTCTACCGGACGGAATTAATCACCGGACGTCGAGATTCAGCTTGTAT 1020  
 QY 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValValProAspGln 360  
 DB 1021 CAGCAACGATGATGCACGATTTCCGGATTACAGAAACTTCGTCGTCGTACCTGACCCAG 1080  
 QY 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380  
 DB 1081 CAAGTCGTTTCAAGCTCCGCGAGATGATCCGCGTGGGTCTCCGGTGGTTTACGACAAG 1140  
 QY 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400

Db	1141	AAACAAGTCGCAAGATTCGGGATTTTAGACAAATACGCCGAAGATTCATCGAACATTAAG	1200
Qy	401	TrpIleAspAlaProAspCysPheHisLeuTrpAsnAlaTrpGluProGlu	420
Db	1201	TGGATTGATGTCGAGATTCGCTTCTGCTTCATCTCTGGAACGCTTGGGAAGACCCAGAA	1260
Qy	421	ThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu	440
Db	1261	ACAGATGAAGTCGTCGATAGGGTCTCTGTAAGTCTCACACAGACTCAATTTTCAACGAG	1320
Qy	441	SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu	460
Db	1321	TCTCAGCAGAACTCAAGAGTGTCTCTGCTGAATTCGCCCTGAATCTCAAAACCGGTGAA	1380
Qy	461	SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet	480
Db	1381	TCAACTCGCCGTCGATCATCTCCACGAGATCAACAAGTCAACCTCGAAGCAGGAGATG	1440
Qy	481	ValAsnArgAsnMetLeuGlyArgGlyThrIlyspPheAlaTyrLeuAlaLeuAlaGluPro	500
Db	1441	GTCACAGAAACATGCTCGGCCGTAAACCAAATTCGCTTACCTTGGCTTTAGCCGAGCCG	1500
Qy	501	TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis	520
Db	1501	TGGCTAAAGTCTCAGGATTCGCTAAGTTGATCTCACTACTGGAGAAGTTAAGAAACAT	1560
Qy	521	LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGlyGlu	540
Db	1561	CTTTACGCGCATACCGTTACGAGGAGAGCCTCTGTTTCTCCCGGAGAGGAGAGAG	1620
Qy	541	GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu	560
Db	1621	GAAGACGAAGATACATCTCTGTTCTGTCACACGAGAGAGACATGAAATCCGAGTTA	1680
Qy	561	GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal	580
Db	1681	CAGATAGTTAACGCGGTAGCTTAGAGGTTGAACCAACGGTTAACTTCGTCAGGGTT	1740
Qy	581	ProTyrGlyPheHisGlyThrPheIleGlyAlaAspLeuAlaLysGlnValVal	599
Db	1741	CCGTACGAGATTCACGGTACATTCATCGGACCGCATGATTTGGCGAAGCAGGTCGTG	1797
RESULT 3			
AB026549			
LOCUS	AB026549	3869 bp	DNA linear
DEFINITION	Arabidopsis thaliana gene for neoxanthin cleavage enzyme, complete cds.		
ACCESSION	AB026549		
VERSION	AB026549.1	GI:16416373	
KEYWORDS	neoxanthin cleavage enzyme.		
SOURCE	Arabidopsis thaliana (chale cresse)		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.		
REFERENCE	1	Iuchi,S., Kobayashi,M. and Shinozaki,K.	
AUTHORS	Characterization of neoxanthin cleavage enzyme from Arabidopsis thaliana		
TITLE	Unpublished		
JOURNAL	2	(bases 1 to 3869)	
AUTHORS	Iuchi,S. and Shinozaki,K.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-APR-1999) Satoshi Iuchi, RIKEN, Plant Mol. Bio.; 3-1-1 Koyuadai, Tsukuba 305-0074, Japan		
	(E-mail:iuchi@rtc.riken.go.jp, Tel:81-298-36-4359)		
FEATURES	Location/Qualifiers		
source	1..3869		
	/organism="Arabidopsis thaliana"		
	/mol_type="genomic DNA"		
	/strain="Col."		
	/db_xref="taxon:3702"		
	1831..3630		
CDS	1831..3630		

/codon_start=1			
/product="neoxanthin cleavage enzyme"			
/protein_id="BAB70609.1"			
/db_xref="GI:16416374"			
/translation="MASFTATAAASGRWLGGNHTPPLSSQSSDLSYCSLPMASRV			
TRLNVSALHPKTAADPSVQIAGNFAPVNEQPVRRNLPVVGLKPDGKGVYVRNGANP			
FLVSHKLPKLPKTAADPSVQIAGNFAPVNEQPVRRNLPVVGLKPDGKGVYVRNGANP			
LHPVTGHHFFDGDGMVHAKVPEHGSAYACRFOTNRNFVQERQLGRPVPPKAIIEUH			
GHTGIARLMLFYARAAAGLVDPAHGTGVANAGLVYFNGLLMSSEDDLPVQVQITENG			
DLTKGRFDPDGMESTIAHPKVDPESEGFALSDVSVKPYLKYTRFRSPDGTKSPD			
VETQLQPTMHDFAITENFVVVDQVVFVKLEPMIRGGSPVYDKNKVAFGLDKY			
AEDSNIKWIDAPDCFLHNAWEEPEDEVVVIGSCMTTPDSINESDENKLSVLS			
EIRLNKLGTESTRPPIISNEDQOVNLEAGMVRNMLGRKTKFAYLALAEWPVKSGFA			
KVDLTGVEYKHLVYDNRVGGELFLPGEEGEEDEGYLLCFVHDEKTKSELQIVNAV			
SLSEVAIVKLSPVPVGFHGTGFIADDLAKQVV"			
BASE COUNT	1155 a	845 c	773 g 1096 t
ORIGIN			
Alignment Scores:			
Pred. No.:	1.99e-219	Length:	3869
Score:	3150.00	Matches:	599
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	8	Gaps:	0
US-09-758-269-6 (1-599) x AB026549 (1-3869)			
Qy	1	MetAlaSerPheThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThr	20
Db	1831	ATGGCTCTTTTCACGGCAACGGCTGGCTTCTGGGAGATGGCTTGGTGGCAATCATACT	1890
Qy	21	GlnProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet	40
Db	1891	CAGCGCCAVTATCTCTCTCAAGCTCCGACTTGAGTTATTGTAGCTCTCTTACCTATG	1950
Qy	41	AlaSerArgValThrArgLysLeuAsnValSerAlaLeuHisThrProProAlaLeu	60
Db	1951	GCAGTCTGTGCACGTAAGCTCAATGTTTCACTCGGCTTACACTCTCTCCAGCTCTT	2010
Qy	61	HisPheProLysGlnSerSerAsnSerProAlaIleValValLysProLysAlaLysGlu	80
Db	2011	CATTTCCTTAAGCAATCATCAAACTCTCCGCCATTGTTGTTAAGCCCAAGCAAGAA	2070
Qy	81	SerAsnThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAla	100
Db	2071	TCCAACACTAAACAGATGAATTTGTTCCAGAGAGCGCGCGGCGGTGGAGCGCGCG	2130
Qy	101	GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer	120
Db	2131	GAGGGTTTCTTGTACCGCAGAGAGCTACACCGCTTCTTAAACGGGTGATCTTAGT	2190
Qy	121	ValGlnIleAlaGlyAsnPheAlaProValAsnGlnProValArgArgAsnLeuPro	140
Db	2191	GTTCCAGATCGCGGAAATTTTGTCTCGGTGAATGAACAGCCCGTCCGGCTAATCTTCG	2250
Qy	141	ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro	160
Db	2251	GTGTCGGAATACTTCCCGATTCCATCAAGAGGATGTATGTGCGCAACGGAGCTTAA	2310
Qy	161	LeuHisGluProValThrGlyHisHisPhePheAspGlyAspGlyMetValHisAlaVal	180
Db	2311	CTTCACGAGCGGTGACAGGTACCACTTCTTCACGAGACGGGTATGGTTCACGCGTC	2370
Qy	181	LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal	200
Db	2371	AAATTCGAACACGGTTTCAGCTAGCTACGCTTCCCGTTTACTCAGACTTAACCGGTT	2430
Qy	201	GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis	220
Db	2431	CAGGAACGTCAATTTGGGTGACCGGTTTCCCCAAAGCCATCGGTGAGCTTCACGGCAC	2490
Qy	221	ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaGlyIleValAspPro	240

Db 2491 ACCGGTATTCGGCGACTCATGCTATTCTACGCCAGAGCTGCAGCCGGTATAGTCGACCCG 2550  
|||||  
Qy 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValThrPheAsnGlyArgLeuLeuAla 260  
|||||  
Db 2551 GCACACGGAACCGGTAGCTAACGCCGCTTTGGTCTATTCTCAATGGCCGGTATTGGCT 2610  
|||||  
Qy 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280  
|||||  
Db 2611 ATGTCCGAGGATGATTACCTTACCAGATTCAGATCACTCCCAATGGAGATTAAACCC 2670  
|||||  
Qy 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300  
|||||  
Db 2671 GTTGTGCGTTTCGATTTTATGATGACAAATAGAAATCACAATGATTCGCCACCACCGAAGTC 2730  
|||||  
Qy 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeu 320  
|||||  
Db 2731 GACCCGGAATCCGGTGAACCTCTTCGTTTAAGCTACGAGCTGCTTCAAGGCTTACCTA 2790  
|||||  
Qy 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340  
|||||  
Db 2791 AAATACCTTCGATTCCTCCGCGACGGAACCTAAATCACCGGACGCTCGAGATTCAGCTTGAT 2850  
|||||  
Qy 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGln 360  
|||||  
Db 2851 CAGCCCAACGATGATGACGATTTCCGATTTACAGAACTTCGTGCTGCTACCTGACCCAG 2910  
|||||  
Qy 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380  
|||||  
Db 2911 CAAGTCGTTTCAAGCTCCGAGATGATCCGCGGTGGGTCTCCGGTGGTTTACGACCAAG 2970  
|||||  
Qy 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerAsnIleLys 400  
|||||  
Db 2971 AACAAAGTCCGAAGATTCGGGATTTAGACAAATACGCCGAAGATTCATCGAACATTAA 3030  
|||||  
Qy 401 TrpIleAspAlaProAspCysPhePheHisLeuTrpAsnAlaTrpGluLeuProGlu 420  
|||||  
Db 3031 TGGATTGATGCTCCAGATTCCTTCCTTCCTCATCTCTGGAACGCTTGGGAAGGCGCAGAA 3090  
|||||  
Qy 421 ThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440  
|||||  
Db 3091 ACAGATGAAGTCGTCGTGATAGGTCCTGTATGACTCCACACAGATCAATTTTCAACGAG 3150  
|||||  
Qy 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460  
|||||  
Db 3151 TCTGACGAGAATCTCAAGAGTCTCTGTCTGAATCCGCTGAAATCTCAAAACCGGTGAA 3210  
|||||  
Qy 461 SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480  
|||||  
Db 3211 TCAACTCCCGCTCCGATCATCTCCAAACGAAGATCAACCAAGTCAACCTCGAAGCAGGGATG 3270  
|||||  
Qy 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500  
|||||  
Db 3271 GTCAACAGAAACATGCTCGCCGCTAAACCAATTCGTTACTTACTGGCTTTAGCCGAGCG 3330  
|||||  
Qy 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis 520  
|||||  
Db 3331 TGGCCTAAAGTCTCAGGATTCGCTAAAGTTGATCTCACTACTGGAGAAGTTAAGAAACAT 3390  
|||||  
Qy 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGlyGlyGlu 540  
|||||  
Db 3391 CTTTACGCGGATTAACCGTTTACGAGGAGAGCCCTCTGTTTCTCCCGGAGAAGGAGAGAG 3450  
|||||  
Qy 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560  
|||||  
Db 3451 GAAGACGAGGATACATCCTCTGTTTCGTTTCACGACGAGACATGGAATTCGAGTTA 3510  
|||||  
Qy 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580  
|||||  
Db 3511 CAGATAGTTAACCGCTTAGCTTAGAGGTTGAAGCAACCGTTAAACTTCCGTCAGGGTT 3570  
|||||  
Qy 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGlnVal 599  
|||||

Db 3571 CCGTACGATTTCCAGGTACATTCTATCGAGCCGATGATTTGGCGAAGCAGGTCGTG 3627  
|||||  
RESULT 4  
AB028617/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
1 (sites)  
Sato,S., Nakamura,Y., Kaneko,T., Kato,T., Asamizu,E. and Tabata,S.  
Structural analysis of Arabidopsis thaliana chromosome 3. I.  
Sequence features of the regions of 4,504,864 bp covered by sixty  
P1 and TAC clones  
DNA Res. 7 (2), 131-135 (2000)  
20277480  
PUBMED  
10819329  
2 (bases 1 to 52232)  
Sato,S., Nakamura,Y., Kaneko,T., Kato,T., Asamizu,E. and Tabata,S.  
Submitted (09-JUN-1999) Yasukazu Nakamura, Kazusa DNA Research  
Institute, Department of Plant Gene Research; 1532-3, Yana,  
Kisarazu, Chiba 292-0812, Japan [E-mail: ynakamu@kazusa.or.jp,  
Tel:81-438-52-3935, Fax:81-438-52-3934]  
Address for correspondence: kaos@kazusa.or.jp  
For the latest information on annotation of this clone, please see  
http://www.kazusa.or.jp/kaos/cgi-bin/agd/graph.cgi?c=MOA2  
Genes with similarity to proteins in the databases are described in  
'product' or 'note' qualifiers. Genes that have no significant  
protein similarity are described as 'unknown protein'.  
The software programs used to predict genes include: Grail  
(Informatics Group, Oak Ridge National Laboratory,  
http://compbio.ornl.gov/Grail-1.3/),  
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),  
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of  
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and  
SplicePredictor (Volker Brendel, Stanford University,  
http://gremlini.zoology.ualb.ca/edu/cgi-bin/sp.cgi).  
Genes encoding tRNAs are predicted by tRNAscan-SE  
(Sean Eddy, Washington University School of Medicine, St. Louis,  
http://genome.wustl.edu/eddy/tRNAscan-SE/).  
This sequence may not be the entire insert of this clone. It may be  
shorter because we remove overlaps between neighboring submissions.  
The 5' clone is MLN21 and the 3' clone is MIE1.  
Location/Qualifiers  
1. 52232  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="MOA2"  
/clone\_lib="Mitsui PI"  
join(1596..2132,2225..2349,2455..2501,2648..2725,  
2807..2931,3021..3087,3216..3278,3370..3439,3529..3629,  
3718..3822,3906..4091)  
/note="gene\_id:MOA2.2"  
/codon\_start=1  
/evidence=not experimental  
/product="glycolate oxidase"  
/protein\_id="BAB01334.1"  
/db\_xref="GI:11994212"  
/translation="MEITNVTEYDAIAKQFLKMYDYVYASGAEDQWTLQENRNARFAR  
ILFPRILIDSVKEIDMTTTLVGLFSIMPIMVAPTMQKNAHPDGEYATARAASAAGTI  
MTLSSWATSSVEEVAATPGPIRFQLYVYKNNRVVQVLRRAERAGFKALITVDTPT  
LGRESIDKNRFTLPNLTAKNFEGLDGLKMDKDEANDSLASYAGQIDRTLSWKDVQV  
LQITKLPILVKVGTEDARIALQAGAGLIIVSNHGARDLYVPATISALEEVVKAT  
QGTIPVFLDGGVRRGTDFVKALALGASGIFIGRPVVFSLAAEGAGVKVQLQMRDEF

CDS	ELTWALSGCRSLKEISRNHITTEWDTPRPSARL" complement (join (4591. .4659,4763. .4924)) /note="gene_id:MOA2.3 unknown protein" /codon_start=1 /evidence=not experimental /protein_id="BAB01335.1" /db_xref="GI:11994213" /translation="MIGRRAGTRNVGMRRDSSLTRFVDSVFYFRFLAEFEILPVLPM LITVYFKDLTSRPEYRNILVEKPGGSDWPF" complement (11870. .13669) /note="gene_id:MOA2.4" /codon_start=1 /evidence=not experimental /product="9-cis-epoxycarotenoid dioxygenase" /protein_id="BAB01336.1" /db_xref="GI:11994214" /translation="MASFTATAVSGRWLGNGHTQPPLSSSSQSDLSYCSLPMASRV TRKNVSSALHTPPALHPFKQSSNSPAIVVKPKAKESNTKQMLFORAAAAALDAAG FLVSHKHLPLPKTADPSVOLAGNFAVNEQPVERNLVVKLPDLSIKGVYVNGNP LHEPVGHHFPGDGMHVAKEFGSASYACFTQNRVQERQLRPVFPKALGELH GHTGALMLFQAASAAGI VDPAHGTGVANAGLVYFNGRLLAESDDLPYQVITPNG DLKTVPDFDQJUESITIAHPKVPDESSELFALSVDVSKPYLKFYRFPDQTSKD VEIQDQPTWHDFAITENFVVPDQVFPKLPKPMIRGGSPVVDKKNVARFGILDKY AEDSNIKWIDPDCFCFLHNAWEEPETDEVVIGSCMTPPDSI FNESDENLKSJLY EIRLNKLTGESTREPIISNEDQOVNLEAGMVRNMLGRKTKFAYALAAEPWPKVSGFA KYDITTEVKKHLGDNRYGGEPLFLPGEGEDEGILCFVHDEKTKWSELQIUNAV SLEVEATVPSPRVPHGTFIGADDAKQVV" join (29980. .30324,30421. .30528,30616. .30660,30841. .30913, 31021. .31133,31239. .31298,31376. .31492,31588. .31649, 31733. .31793) /note="contains similarity to RNA-binding protein gene_id:MOA2.5" /codon_start=1 /evidence=not experimental /protein_id="BAB01337.1" /db_xref="GI:11994215" /translation="MAVTBEMPTDVVVVDVKEVSTSEKILDEIGIEKSSITDSKTE ESRUDMKLVAMFKKINPLAKEPPSYDYPKKNQVAKANQFLPADDFETIKKSGEE FDLDAKKDNLTKRRNYSQGRRLTGRIKSAQREDSIRRTVYSDIDQSVTEBGLAGL FSNCGVQVDCRICDGHPSVTRFAFVEADQGAHEALSLGTMFGFVPRVLPSTKTAI LPVNPFTLPRSEDEMCRTIYCTNIDKKVQADVRNFFESACGEVTRLLRLDQLH STRIAFVEFALADSALNCSGMVGSQPIRVSPSKTPVRITRPPSTN" complement (32182. .36456) /note="gene_id:MOA2.6" /codon_start=1 /evidence=not experimental /product="disease resistance complex protein" /protein_id="BAB01338.1" /db_xref="GI:11994216" /translation="MANSYLSNCANVWVERINTSQEVLVELCKGSSALLKRLKVALY TANPLADAQRAEHRVRLWLTGIDAFQAEIDIDELQTEALRRRVVAEAGGLGG LFQNLMAGRAIQKKEIPKMEKVVRLEHVKHTEVIGLKEYSETREPQWRQSRSP DDLPGQLGRVEDKALVLLSDDBSISCKPAVISVWGMPPGVGKTLTLEIVFNDR VTEHFEVNIISAGINFTVTKAVLQDITSSAVNTDLPSIQOLKTKLTKGRFLL VLDFWMSSEWSESFOVATDAEGSKI VLTTRSELSVTAKAEKIYQMKLMTNEC WELLSRAPGNISVGSINQELGIGKRIAQCKGLPAAARAIAHLSKSNPDNDWYAV SKNFSSTNLSVLKLSYDSLPPQLKRCFCALCSIFPKGHVDFREELVLLWMAIDLLEY QPRSRLEDIGNDGLVQAQFFQLDITMTSFVMDLMDLAKAVSGDFCFRLDD DNIPEITSTRHFSQSCDQSAVFRSICGAEPRLTILPNSPTLSLESQLTEKVLN PLNALSGRLTSLSHVQIITNLPSKLGKLLRYLDLSSTKIKELPEFVCTLGNLT LLSNCRDLSLPKSIABLNLRLDLVGPLVEMPPGIIKLSRLQKLSNVEIGLSGA GHELEKSHURGTLRSELQNVAFSAEKADGLKRPFLDGLLKLWTVKXSGVFNCS FNALACQKEHURLMLEPHPLKTCIESYQGAAPPKWIGDSFFGITSVLSNCLCI SLPVGQLPSLKYSILFNILQKVLGDFFGENNSRGVFPQSILQILKFGMPRWDM ICPLEDGIIFPCKLIQRCPSLRKFPGLPSSTEVITSDCPLRAVSGSGSFRS LTNIPESPAITPSMRSELSSPTGNPKSDASTSAQGFASSQSNDNVEVTSLSLPS LPKDRQTFDEQVTOGLSLPQOFEPAVLSARYSGVLSIDIPSLPYMKTSLVPP KNESILPAGSSSYOYHQYKSSVPSRSEAIKPSQDDEIDMEYKVDIDSHLME LPQNLQSLHSDCSGLTSLENLTESYPNLHELLIIACHLESFPGSHPTTKTLXI RDKCKLNFTLSQTRSYQLEYLFIGSSCNLNVLPFLPKRLSIRDCSPFTK SIHAGLDGRIALRESIRDCNLETFPQGLPTPKLSMLLSNCKKLQALPKFLGL TSLSLSLFIKPEIETIPGGGFPNSNLTLCISLCKLTPRIEWGLRDLNLRNLIEDG
CDS	GNEIDIRSPBEGLPKSVFSLSRISRPENLKNLRKPHDTKAETMEISGCDKLQIS1 DEDLPPLSLCLIRISCSLLTETFAEVETEFEFFKVLNIPYVEIDGRIFS" 38132. .41296 /note="gene_id:MOA2.7" /codon_start=1 /evidence=not experimental /product="disease resistance complex protein" /protein_id="BAB01339.1" /db_xref="GI:11994217" /translation="MTGIGEMFLAAFLQALFQTLVSEPPFRSFFKRRLENLENLERLST ALHTITAVLIDAEKQITNEVBEKWNELRDVVYHABDADDIATEALRINI GAESS SNRLQRLGMSGLDFLDGNSHLETRLEKVTIRLERLASQRIKELTAMIPKOR LPTTSLVDESEVFGRDDKDEIMRFLIPENGKNGITVAIVIGIVGKGKTLTQLLYN DOHVSFPGTKVAHYSEEDVFKITKVYVESVSRPCESTDLDLVOLKELRTGNG LPFILLVDDLWNEFNADWDLRQFFTHAAGSQSLVITRQRVASIMCAVHNHLQPL SDGCSWLSFMKTVFGNQPCLNREIGDLAERI VHCKRGLPLAVKTLGGVLFEGKVI E WAEVLSRINWDLPADKNLPLVRSVYYPALHKRCFCAYCSI FPKGHAPEKOKVLL MKEGFLQOITRSSKNLEELGNEYFSELESLLQKTKRTYIMHDFINELAQFASGEFS SKYFDGCKLOVSRTRYLSYLRDNYAEPMEFALREVKFLRTPLSLTNSSSCCLD QMVSEKLLPTLRLVLVLSHYKIARLPDPFKNI SHARFLDLSRTELEKLPKSLCYM YNIQTLLSVCSSELPTDISNLINRYLDLIGTKLRQMRPRFGRKLSQTLTTFV SASDGRSISGLGDLHLGKLI VEQRVVDVADAANLNSKKHLREIDFVWKTGSS SSNNTPHRTONEAEVFEKPHRIEKLAIERYKGRFPDMLSDREFIVICIRL ECQYCTSLPSLGQLPCLKEHLISGMVGLQSIGRKFYFSDOQLRDQDOQPSLETLPF DNLPDQEWLDVVRTRGDLPPSLKLPILRCPELTGTLPTLPPLSLI SLHIYKGLLDF OPDHEYSYVENLOTLSIKSSCDTLVAFPLNHEANLKLVDQCTSLYSELSEHLRG PNALNRLINCNQLQLPKLNALPQNLOVTITNCRYLRQMEQOQPYHHQPHLPRS NVSGSPKSHGSHSIDSRSRYD" 42345. .42872 /note="gene_id:MOA2.8 unknown protein" /codon_start=1 /evidence=not experimental /protein_id="BAB01340.1" /db_xref="GI:11994218" /translation="MSDLVIALVAFVLFVVFILSCIMGKEDVHSPPLPRGQLR MPYVRKDKDLGGDGTSAISVAGTDIYGGSGVGGHGRGIITSDTKLPSPPPSPUP PRVQRPKSVSRUKDBELIDTGTAMLVASSLLTSGSASSCGSGSHGCGGGGGGG GGGGGGGGGGGGGG" complement (join (42980. .43465,43561. .43668,43952. .44119, 44215. .44460,44539. .44677,44759. .44977,45060. .45754)) /note="gene_id:MOA2.9" /codon_start=1 /evidence=not experimental /product="terpene synthase-like protein" /protein_id="BAB01341.1" /db_xref="GI:11994219" /translation="MKPKVRMDLKSNNSDNERIRLTHLMLNLGIAYHFEIDEILG QAQNLDDIIAKENDLETISTWFEVRLRGYMPCKSIGDLSHICHIYNLAFDFT VDAPNFKGDRGFKESLAEDIRGMQLVEAAHLGTPSDIMDEALSFTYRLESNTS NHTATASPHLSKHQIALYRARNLEILVAREYISFEQEEDHDETLKFAKLNFY CQHYIQELKDLTKWMLDASKLPYRDIRIVIEVIFGALALIFEPYSIGRIIVTKI QWTVFNDTCDAYGLTEVTSVDSVDFQRWDLGDI EKLPSYKIVFVSGVTEEBEQ EMTPQGRSRIYQVAVDEIKKLKAYLAI SKWARASHVPTFEYMEFMQMTSDHFAAY SFTAGSDCNQTCWYKPKWAEALNGVRIKNDINTPEQMSRGEVAKNGYMK OHCVSEKAAIGEMNKIYSNYKIIIMEYLVTAVPREILETFDNLNSKPLI IKTLLTM FRLVNDGVTEVHLKLTSELSRGEVANGLDYFMKQGVTKEEASQELRKNKNYKVY MEEFMTHDLHPQVFLRCHNIARIIDFVTEVDGYCDPKGKIENFMISLYLHPINTP
CDS	Alignment Scores: Pred. No.: 7.97e-218 Length: 52232 Score: 3150.00 Matches: 599 Percent Similarity: 100.00% Conservative: 0 Best Local Similarity: 100.00% Mismatches: 0 Query Match: 100.00% Indels: 0 DB: 8 Gaps: 0 US-09-758-269-6 (1-599) x AB028617 (1-52232)
Qy	1 MetAlaSerPheThrAlaThrAlaAlaValSerClyArgTrpLeuGlyGlyAsnHisThr 20 
Db	13669 ATGGCTTCTTTCACGGCAACGGCTGCGGTTTCTGGGAGATGGCTTGGTGGCAATCATCAT 13610 



QY 21 GlnProProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet 40  
DB 13609 CAGCGCCCAATATCGCTTCTCAAGCTCCGACTTGAGTTATTTAGCTCTTACCTATG 13550  
QY 41 AlaSerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProProAlaLeu 60  
DB 13549 GCCAGTCGTGCACAGTAAGCTCAATGTTTCATCTGCGCTTCACTCTCTCAGCTCTT 13490  
QY 61 HisPheProLysGlnSerSerSerProAlaIleValValProLysAlaLysGlu 80  
DB 13489 CATTTCCCTTAAGCAATCATCAAACTCTCCCGCAITTTGTTAAGCCCAAGCAAGAA 13430  
QY 81 SerAsnThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaAlaAlaAla 100  
DB 13429 TCCAACACTAAACAGATGAATTTGTTCCAGAGAGCGCGCGAGCTGTGACCGCGG 13370  
QY 101 GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer 120  
DB 13369 GAGGGTTCTTGTGAGCCACAGAGAGCTACACCGCTCTCTTAAACGGCTGATCCTAGT 13310  
QY 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgAsnLeuPro 140  
DB 13309 GTTCAGATCGCCGGAATTTGCTCCGGTGAATGAACAGCCGCTCGGCGTAATCTTCG 13250  
QY 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160  
DB 13249 GTGGTCGGAAAACTTCCCGATTCCATCAAGAGAGTGTATGTGCGCAACGGAGCTAACCCA 13190  
QY 161 LeuHisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaVal 180  
DB 13189 CTTACGAGCGGTGACAGGTACACCTTCTTCGAGGAGAGCGGTATGTTACCGCGTC 13130  
QY 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200  
DB 13129 AAATTCGAACACGGTTACGTAGCTACGCTCTCCCGGTTTACTCACACTAACCGGTTGTT 13070  
QY 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220  
DB 13069 CAGGAACGTCAATGGGTGCGACCGGTTTTCCCAAAGCCATCGGTGAGCTCACCGCCAC 13010  
QY 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspPro 240  
DB 13009 ACCGGTATGCCGACTCATGCTATTTCACGCCAGAGCTGACGCCGGTATGTCAGCCCG 12950  
QY 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260  
DB 12949 GCACACGGAACCGGTAGTAAACCGGTTTGGTCTATTCAATGGCGCGTTATTGCT 12890  
QY 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280  
DB 12889 ATGTCGGAGGATGATTTTACCTTACCAGGTTCCAGATCACTCCCAATGGAGATTTAAAAACC 12830  
QY 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300  
DB 12829 GTTGTCCGTTCCGTTGATGGCAATAGAAATCCAAATGATGCCACCCCGAAAGTC 12770  
QY 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysPheProTyrLeu 320  
DB 12769 GACCCGGAATCCGGTGAATCTTCGTTTAACTAGCTACGAGCTCGTTTCAAGCCTTACCTA 12710  
QY 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340  
DB 12709 AAATACTTCCGATTTCTACCGGACGGAACCTAAATCACCGGACGTCGAGATTCAGCTTAT 12650  
QY 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValValProAspGln 360  
DB 12649 CAGCAACGATGATGACGATTTCCGATTTACAGAGAACTTCGTCGTCTGCTACCTGACCCAG 12590  
QY 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValValTyrAspLys 380  
DB 12589 CAAGTCGTTTCAAGCTCCCGAGATGATCGCGGTGGGTCTCCGGTGGTTTACGACAG 12530  
QY 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400

DB 12529 AACAGGTTCGAAGATTTCGGATTTTAGACAAATACGCCGAAGATTTCATCGAACATTAA 12470  
QY 401 TrpIleAspAlaProAspCysPheHisLeuTyrAsnAlaTyrGluGluProGlu 420  
DB 12469 TGGATTGATGTCACAGATTGCTTTCGCTTCCATCTCTGGAAGCGTTGGGAAGAGCCAGAA 12410  
QY 421 ThrAspGluValValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440  
DB 12409 ACAGATGAAGTCGTCGTATAGGGTCTGTATGACTCCACAGACTCAATTTTCAACGAG 12350  
QY 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460  
DB 12349 TCTCAGCAGAAATCTCAGAGTGTCTGTCTGAAATTCGCTCTCAAAACCGGTGAA 12290  
QY 461 SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480  
DB 12289 TCACTCGCCGTCGATCATCTCCACGGAAGATCAACAACTCACTCGAGCAGGATG 12230  
QY 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500  
DB 12229 GTCAACAGAAACATGCTCGCGGTAAACCAAAATTCGCTTACTTGGCTTTAGCCGAGCG 12170  
QY 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis 520  
DB 12169 TGGCTTAAGTCTCAGGATTCGTAAGTTGATCTCTACTTGGAGAAAGTTAAGAAACAT 12110  
QY 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGlyGlu 540  
DB 12109 CTTTACGCGATAACCGTTACGAGGAGAGCTCTGTTTCTCCCGAAGAGGAGAGAG 12050  
QY 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTyrLysSerGluLeu 560  
DB 12049 GAAGACGAAGATACATCTCTGTTTCGTTCCACGAGAGACATGGAATTCGAGTTA 11990  
QY 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580  
DB 11989 CAGATAGTTAACGCGGTAGCTTAGAGGTTGAAGCAACGTTAAACTTCCGTCAAGGTT 11930  
QY 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspLeuAlaLysGluValVal 599  
DB 11929 CCGTACGATTTACGGTACATTCATCGAGCCGATGATTTGGCGAAGCAGGTCGTG 11873

## RESULT 5

AY056255

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 2331)

AUTHORS

Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Ban, J., Chung, M.K.,

Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,

Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,

Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,

Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,

Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,

Davis, R.W., Ecker, J.R. and Theologis, A.

Arabidopsis Full Length cDNA Clones

Unpublished

2 (bases 1 to 2331)

Yamada, K., Ban, J., Banno, F., Dale, J.M., Goldsmith, A.D., Lee, J.M.,

Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y.,

Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R.,

Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,

Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C.,

Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,

Satou,M., Seki,M., Shinn,P., Southwick,A., Tracy,S.E.,  
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.  
Direct Submission  
Submitted (12-SEP-2001) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J.,  
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,  
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the  
sequencing and annotation of the RAFL cDNAs: Yamada,K., Banno,J.,  
Banno,F., Dale,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S.,  
Quach,H.L., Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S.,  
Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C.,  
Koesema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M.,  
Palm,C.J., Shinn,P., Southwick,A., Tracy,S.E., Davis,R.W.,  
Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to  
this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)  
contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis  
genome submitted to GenBank.

#### FEATURES

source  
Location/Qualifiers  
1. .2331  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="RAFL08-11-Hi6 (R11094)"  
/note="This clone is in a modified pBluescript vector  
(FLC-1) as a BamHI/XhoI insert."  
ecotype: Columbia  
1. .2331  
/gene="At3g14440"  
1. .122  
/gene="At3g14440"  
123. .1922  
/gene="At3g14440"  
/codon\_start=1  
/evidence=experimental  
/product="putative 9-cis-epoxycarotenoid dioxygenase"  
/protein\_id="AA107104.1"  
/translation="MASTATAAVSGRWLGNGHTQPPPLSSSQSDLSYCSLPMASRV  
TRKLVNSALHTPPALHPKQSSNSPAIVVKPKAKESNTKQMLFORAAALDAARG  
FLVSHKHLPLPKTADPSVOLAGNFAPVNEQVPRNLVVGKLPDSIKGVYVRGANGP  
LHEPVTHHFPDGGMVAHKFEHGSASYACRFQTRNFVQERQLRPVFPKALGELH  
GHTGIARLMUFYAAAGMI VDPAHGTGVANAGLVYFNGRLAMSEDDLPYQVQITPNG  
DLKTVGRFDGQJLESTMIAPKVPDESSELFSYDVSKPYLKYPFRFPDGTGKSPD  
VEIQLDQPTWMHDFAITENFVVDPQOVFKLPEMILGSGSPVVDKKNKVARFGILDKY  
AEDSSNTKIDAPCFCHLWNEEPETDEVVVGSCMTTPDPSIFNESDENLKSVL  
EIRLNLTGSTRPIISNEDQVNLKAGMWRMLGRKTKFAYLALAEWPVKVSGFA  
KVDLTTGEVRKHLGDMRYGGEPUFLFEGGEDEYILCFVHDEKTKWSELQILVNAV  
SLEAVATVKUPSRYPYFGHTFGADLAKQVV"  
misc\_difference  
1234  
/gene="At3g14440"  
/note="compared to genomic sequence resulting in an amino  
acid sequence difference"  
/replace="g"  
1923. .2331  
/gene="At3g14440"  
misc\_difference  
2059  
/gene="At3g14440"  
/note="not present in genomic sequence"  
misc\_difference  
2312  
/gene="At3g14440"  
/note="not present in genomic sequence"  
misc\_difference  
2315  
/gene="At3g14440"

BASE COUNT 615 a 560 c 533 g 623 t  
ORIGIN  
/note="not present in genomic sequence"

#### Alignment Scores:

Pred. No.: 3.14e-219 Length: 2331  
Score: 3143.00 Matches: 598  
Percent Similarity: 99.83% Conservative: 0  
Best Local Similarity: 99.83% Mismatches: 1  
Query Match: 99.78% Indels: 0  
DB: 8 Gaps: 0

US-09-758-269-6 (1-599) x AY056255 (1-2331)

Qy 1 MetAlaserPheThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThr 20  
Db 123 ATGGCTCTTTTCACGGCAACGGCTGGGTTTCTGGGAGATGGCTTGGTGGCAATCATACT 182  
Qy 21 GlnProProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet 40  
Db 183 CAGCCGCCATTATCGTCTTCTCAAAGCTCCGACTTGAGTTATTGTAGCTCTCTTACCTATG 242  
Qy 41 AlaserArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProProAlaLeu 60  
Db 243 GCCAGTCGTCTCACACGTAAGCTCAATGTTTCAATCTCGCTTCACACTCTCCAGCTCTT 302  
Qy 61 HisPheProLysGlnSerSerAsnSerProAlaIleValLysProLysAlaLysGlu 80  
Db 303 CATTTCCCTAAGCAATCATCAAACTCTCCGCCATTGTTTGAAGCCCAAGCCAAAGAA 362  
Qy 81 SerAsnThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaAlaLeuAspAlaAla 100  
Db 363 TCCACACACTAAACAGATGAATTTGTTCCAGAGAGCGCGCGGAGGTTGGACCGCGCG 422  
Qy 101 GluGlyPheLeuValSerHisGlyLysLeuHisProLeuProLysThrAlaAspProSer 120  
Db 423 GAGGGTTTCTTGTGCCACACGAGAGACTACACCGCTTCTCTAAACGGCTGATCTCTAGT 482  
Qy 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGlnProValArgAsnLeuPro 140  
Db 483 GTTCAGATCCCGGAAATTTTGCTCCGGTGAATGAACAGACCGCGCTCCGGCGTAACTTCCG 542  
Qy 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160  
Db 543 GTGGTCGAAAACCTCCCGATTCCATCAAGAGAGTGTATGTGCGCAACGAGCTAACCCA 602  
Qy 161 LeuHisGluProValThrGlyHisHisPhePheAspGlyAspGlyMetValHisAlaVal 180  
Db 603 CTTTCCAGCGCGGTGACAGGTCACTTCTTCGACGGAGACGGTATGGTTTACGCGCTC 662  
Qy 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200  
Db 663 AAATTCGAACACGGTTCAGCTAGCTACGCTTGCCTGCGGTTTACTCAGACTAACCGGTTTGT 722  
Qy 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220  
Db 723 CAGNAGCTCAATTTGGTGCACCGGTTTCCCAAGCCATCGGTGAGCTTTCACGGCCAC 782  
Qy 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspPro 240  
Db 783 ACCGGTATTGCCGACTCATGCTATTCTACGCCAGAGCTGCAGCCGCTATAGTCGACCCG 842  
Qy 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260  
Db 843 GCACACGGAACCGGTGTAGCTAACCGCGGTTTGGCTTATTTCATATGCGCGGTATTGGCT 902  
Qy 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280  
Db 903 ATGTCGGAGGATGATTACCTTACCAAGTTCCAGATCACTCCCAATGGAGATTTAAARACC 962  
Qy 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300  
Db 963 GTTGGTCGGTTTCGATTTTGTATGGACAATTAGAAATCCCAATGATTGTCGCCACCCGAAAGTC 1022

QY 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeu 320  
|||  
Db 1023 GACCCGAATCGGTAACCTCTCGCTTTAAGCTACGCTCGTTTCAAAGCCTTACCTA 1082  
|||  
QY 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluLeuLeuAsp 340  
|||  
Db 1083 AAATACTTCGGATTCTACCGGACGAACTAATACCGGAGCTCGAGATTTCAGCTTGAT 1142  
|||  
QY 341 GlnProThrMetHisAspPheAlaLeuThrGluAsnPheValValProAspGln 360  
|||  
Db 1143 CAGCCAAACGATGATGACGATTCGCGATTACAGAGAACTTCGTCGTCTACCTGACCCAG 1202  
|||  
QY 361 GlnValValPheLysLeuProGluMetLleArgGlyGlySerProValValTyrAspLys 380  
|||  
Db 1203 CAAGTCGTTTCAAGTCGCGAGATGATCCTCGGTGGTCTCCGCTGGTTCACGACAG 1262  
|||  
QY 381 AsnLysValAlaArgPheGlyLleLeuAspLysTyrAlaGluAspSerSerAsnLys 400  
|||  
Db 1263 AACAGGTCGCAAGATTCCGGATTATAGCAATACGCGAAGATTCTCGAACAATTAAG 1322  
|||  
QY 401 TrpLleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluProGlu 420  
|||  
Db 1323 TGGATTGATCTCCAGATTGCTTCTGCTTCATCTCTGGAACGCTTGGAGAGCCAGNA 1382  
|||  
QY 421 ThrAspGluValValLleGlySerCysMetThrProProAspSerLlePheAsnGlu 440  
|||  
Db 1383 ACAGATGAAGTCGTCGATAGGTCCTGTATGACTCCACACAGACTCAATTTTCAACGAG 1442  
|||  
QY 441 SerAspGluAsnLeuLysSerValLeuSerGluLleArgLeuAsnLeuLysThrGlyGlu 460  
|||  
Db 1443 CTCGACGAGAACTCTCAAGAGTGTCTGTCTGAAATCCGCTCAATCTCAAAACCGGTGAA 1502  
|||  
QY 461 SerThrArgArgProLleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480  
|||  
Db 1503 TCAACTCGCGTCCGATCATCTCCACGAGAGTCAACAGTCAACTCGAAGCAGGATG 1562  
|||  
QY 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500  
|||  
Db 1563 GTCAACAGAAACATGCTCGCGCTAAACCAAATTCGCTTACTTGGCTTTAGCCGAGCGG 1622  
|||  
QY 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrGlyGluValLysLysHis 520  
|||  
Db 1623 TGGCTAAAGTCTCAGGATTCCTAAAGTTGATCTCACTACTCGGAGAGTTAAGAAACAT 1682  
|||  
QY 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGlyGlu 540  
|||  
Db 1683 CTTTACGGGATACCGTTACGAGAGAGCGCTCTGTTTCTCCCGGAGAGAGAGAG 1742  
|||  
QY 541 GluAspGluGlyTyrLleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560  
|||  
Db 1743 GAAGACGAAGGATACATCTCTGTTTCTGTTTACGAGAGAGACATGGAATCGGAGTTA 1802  
|||  
QY 561 GlnLleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580  
|||  
Db 1803 CAGATAGTTAACCGCTTAGCTTAGAGTTGAAGCAACGGTTTAAACTTCCGCTCAAGGTT 1862  
|||  
QY 581 ProTyrGlyPheHisGlyThrPheLleGlyAlaAspAspLeuAlaLysGlnValVal 599  
|||  
Db 1863 CCGTACGAGTTTACGGTACATTCATCGGAGCGGATGATTGGCGAAGCAGGTCTGTG 1919  
|||

RESULT 6  
STU276244  
LOCUS  
DEFINITION Solanum tuberosum mRNA for putative 9-cis-epoxycarotenoid  
dioxygenase (nced1 gene).  
ACCESSION  
VERSION AJ276244.1 GI:7209268  
KEYWORDS 9-cis-epoxycarotenoid dioxygenase; nced1 gene.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE  
AUTHORS Burbridge, A., Taylor, I.B. and Thompson, A.  
TITLE Potato putative 9-cis-epoxycarotenoid dioxygenase 1 cDNA  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2164)  
AUTHORS Burbridge, A.  
TITLE Direct Submission  
JOURNAL Submitted (02-MAR-2000) Burbridge A., Plant Science Division, The  
University of Nottingham, Sutton Bonington Campus, Loughborough,  
Leicestershire, LE12 5RD, UNITED KINGDOM  
FEATURES  
source  
1..2164  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4113"  
1..2164  
/gene="nced1"  
73..1887  
/gene="nced1"  
/function="cleavage of 9-cis-epoxycarotenoids"  
/codon\_start=1  
/product="putative 9-cis-epoxycarotenoid dioxygenase"  
/protein\_id="CAB76920.1"  
/db\_xref="GI:7209269"  
/translation="MATTTSHTATNTWIKPLSPSPSKFSGFASNSISLLKQHNQRL  
NINSLOAPPILHPKQSSNYQTKTSTISHPKQNNSSSSISKNVLQVKAAMALD  
AVEGALTGHELEHPLKPTADPRVOISGNFAPVPENPVCQSLPVTGKI PKCVGVVVRN  
GANLPFEPTAGRHFDGDMVHAFQKNGSASACRFTETEFVQEKALGRVPFKAI  
GELHSGIARLMLFYARGLGLDHSRGTVANGLVYFNNRLLMASEDDLPHYKV  
TPTGDLKTEGRFDFDQKLSMTIAHPKLDVPSGELFALSVDVIQPKYLKRFPSKNGE  
KSNDEIPEVEDPTMMHDFAITKFIIPDQVVFMSMIRGSSPVVYDKNKSFRGI  
LDKAKDGSDLKWEVDFPCFHLWNAAWEEPEDEIIVVIGSCMTDPDSIFNECDSEGLK  
SVLSIRLNKTKGSTRKAI IENPDEOVLEAGMVRNKLGRKTQYAVLATAEPKPV  
SGFAKVDLFTGEVEKFLGDNKYGSEPLFLPRDPSKEEDDGYILAFVHDEKEWTSEL  
QIVNATLKLEATVKLPSRVPGFHGTFINANDLANQA"  
BASE COUNT 667 a 437 c 433 g 627 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,74e-157 Length: 2164  
Score: 2291.00 Matches: 433  
Percent Similarity: 81.97% Conservative: 67  
Best Local Similarity: 70.98% Mismatches: 92  
Query Match: 72.73% Indels: 18  
DB: 8 Gaps: 5  
US-09-758-269-6 (1-599) x STU276244 (1-2164)

QY 1 MetAlaSerPheThrAlaThrAlaAlaValSerGlyArgTyrLeuGlyGlyAsnHisThr 20  
|||  
Db 67 GTAGCTATGGCCACTACTTTCACATGCTACAAATACATGATT----- 111  
|||  
QY 21 GlnProProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSer-----Ser 37  
|||  
Db 112 AAGCTAAGTTGTCAATGCCATCATCAAAAGAGTTTGGTTTTCATCAAACTCTATTCT 171  
|||  
QY 38 LeuProMetAlaSerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrPro 57  
|||  
Db 172 CTACTCAAAAATCAACATAATAGACAAAGTCTCAACATTAACTCTCTCTCTCAAGCTCCA 231  
|||  
QY 58 ProAlaLeuHisPheProLysGlnSerSerAsn-----SerProAlaIleVal 73  
|||  
Db 232 CCTATTCTTCATTTTCTTAACAATCTTCAATATATCAACACCAACCACTAGTACAATT 291  
|||  
QY 74 ValLysProLysAlaLysGluSerAsnThr-----LysGlnMetAsnLeuPhe 89  
|||  
Db 292 TCACACCCCAAAACAAGAAAACAACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 351  
|||  
QY 90 GlnArgAlaAlaAlaAlaLeuAspAlaAlaGluGlyPheLeuValSerHisGluLys 109  
|||  
Db 352 CAGAAAGCAGCAGCAATGGCTTTAGATGCTGTAGAAGTGTCTTTAAACAAAACACCACTC 411  
|||

Asteridae; lamids; Solanales; Solanaceae; Solanum.

```
Qy 110 LeuHisProLeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaPro 129
Db 412 GAGCACCCTTTGCCGAAACAGCCAGCCAGAGTCAGATTTCTGGGAATTTTCGCTCCG 471
Qy 130 ValAsnGluGlnProValArgAsnLeuProValValGlyLysLeuProAspSerIle 149
Db 472 GTACCGGAACACCAGTTTGGCAATCTCTTCGGGTACCAGGAAATAATACCCAAATGTGT 531
Qy 150 LysGlyValTyrValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHisHis 169
Db 532 CAAGCGGTTTACGTTGCAAAACAGGAGCTAACCCCTCTTTTCGAACCAACCGCGGACGCCAT 591
Qy 170 PhePheAspGlyAspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyr 189
Db 592 TTCTTCGACGCGAGGTATGGTTTACGCCGTTCAATTCAAATAATGGGTAGCTAGTTTAC 651
Qy 190 AlaCysArgPheThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProVal 209
Db 652 GCTTGGCGGTTTCACTGAAACAGAGAGGTTTGTTCGAAGAAAAGCTTTGGGTGCGCCTGTT 711
Qy 210 PheProLysAlaIleGlyLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPhe 229
Db 712 TTCCTTAAGCCATTGGTGAATACATGCTCACTCTGGAATTCGAAGCTTTATGCTGTT 771
Qy 230 TyrAlaArgAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsnAla 249
Db 772 TACGCTCTGGGCTCTTCGGACTTATTGATCACAGTAGAGGAAGTGGAGTTGCAACGCC 831
Qy 250 GlyLeuValTyrPheAsnGlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGln 269
Db 832 GGTATTAGTCTATTTCATTAACCGATTACTTCTGATGCTGAAGATGATTTGCTTACCAT 891
Qy 270 ValGlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAspPheAspGlyGln 289
Db 892 GTAAAGGTAAACACCCCGCGGATCTTAAAAACAGAGGGTCGATTTCGATTTTCGACGGCCAG 951
Qy 290 LeuGluSerThrMetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAla 309
Db 952 CTAAATCCCATGATAGCTCACCAAGCTCACCCAGCTTCGGTGAGCTATTGCT 1011
Qy 310 LeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGly 329
Db 1012 CTAGCTACGATGATGATTCAGAGCCATACCTCAAGTATTTTCAGATTTTCAAAAATGGG 1071
Qy 330 ThrLysSerProAspValGluIleGlnLeuAspGlnProThrMetMetHisAspPheAla 349
Db 1072 GAAAAATCAAAATGATGTTGAAATCCCATGTTGAAGACCCCAACCAATGATGATTTCCGCA 1131
Qy 350 IleThrGluAsnPheValValValProAspGlnGlnValValPheLysLeuProGluMet 369
Db 1132 ATTACTGAAAAATTCGTCATCCTGATCAACAAAGTCGTTTTCAAGATGCTCGAATG 1191
Qy 370 IleArgGlyGlySerProValValTyrAspLysAsnLysValAlaArgPheGlyIleLeu 389
Db 1192 ATCCGTGGAGGTTCACCGGTGTTTACGACAAGAACAAAGTGTCCCGATTTGGGATCTG 1251
Qy 390 AspLysTyrAlaGluAspSerSerAsnIleLysTrpIleAspAlaProAspCysPheCys 409
Db 1252 GATAAGTACGCGAAGATGGGTCTGATTTGAAATGGGTGGAAGTTTCCTGATTTGTTCTGT 1311
Qy 410 PheHisLeuTrpAsnAlaTrpGluGluProGluThrAspGluValValIleGlySer 429
Db 1312 TTCCATCTCTGGAATGCTTTGGGAAGAACCAAGAACAGATGAATGTTGTTAATTTGGTTCA 1371
Qy 430 CysMetThrProProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerValLeu 449
Db 1372 TGTATGACACACCCAGACTCCATTTTCAATGAATGATGATGAGGACTAAAGAAGCGTTTA 1431
Qy 450 SerGluIleArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsn 469
Db 1432 TCCGAAATTCGCTCTCAATTTGAAAAACAGGGAATCAACAAAGAAAGCCATTAATCGAAAC 1491
Qy 470 GluAspGlnGlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLys 489
```

```
Db 1492 CCGGATCAACAAGTGAATTTAGAGGCTGGAATGGTGAACACAGAACTTGGAGGAA 1551
Qy 490 ThrLysPheAlaTyrIleuAlaLeuAlaGluProTrpProLysValSerGlyPheAlaLys 509
Db 1552 ACACAGTATGCTTATTGGCTATTGCTGAACCATGGCCAAAAGTTTCTGCTTTTGCAAA 1611
Qy 510 ValAspLeuThrThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGly 529
Db 1612 GTAGACCTGTTTACAGGTGAAGTTGAGAAATTCATTATGTTGTCACAAATATGGTGGG 1671
Qy 530 GluProLeuPheLeuProGlyGlu-----GlyGlyGluGluAspGluGlyTyrIleLeu 547
Db 1672 GAGCCTCTTTTACCAAGACCCCAACAGGAGGAGGAGGACGATGTTATATTTTA 1731
Qy 548 CysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaValSer 567
Db 1732 GCTTTCGTTTCATGATGAGAAAGATGACATCAGAGCTTCAAAATTTGTTAATGCAATGACT 1791
Qy 568 LeuGluValGluAlaThrValLysLeuProSerArgValProTyrGlyPheHisGlyThr 587
Db 1792 TTAAGTTAGAGGCAACAGTGAAACCTTCATCAAGAGTTCCTTATGTTTTCATGGAACT 1851
Qy 588 PheIleGlyAlaAspAspLeuAlaLysGln 597
Db 1852 TTCAATAATGCAATGATTTGGCAATCAG 1981

RESULT 7
LOCUS LENEOKANT 2171 bp mRNA linear PLN 22-MAY-1998
DEFINITION Lycopersicon esculentum mRNA for nine-cis-epoxycarotenoid
dixygenase.
ACCESSION 297215
VERSION 297215.1 GI:2769641
KEYWORDS nine-cis-epoxycarotenoid dioxygenase.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; lamids; Solanales; Solanales; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE
1 Burbridge, A., Grieve, T. M., Jackson, A., Thompson, A. and Taylor, I. B.
Structure and expression of a cDNA encoding a putative neoxanthin
cleavage enzyme (NCE) isolated from a wilt-related tomato
(Lycopersicon esculentum Mill.) library
J. Exp. Bot. 47, 2111-2112 (1997)
2 Burbridge, A.
Direct Submission
Submitted (01-JUL-1997) Burbridge A., The University of Nottingham,
Physiology and Environmental Science, Sutton Bonington Campus,
Loughborough, Leicestershire, LE12 5RD, UK
Revised by [3]
3 (bases 1 to 2171)
Direct Submission
Burbridge, A.
Submitted (09-JAN-1998) Burbridge A., The University of Nottingham,
Physiology and Environmental Science, Sutton Bonington Campus,
Loughborough, Leicestershire, LE12 5RD, UK
On Jan 13, 1998 this sequence version replaced gi:2243153.
COMMENT Location/Qualifiers
FEATURES
source 1..2171
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
79..1896
/codon_start=1
/product="nine-cis-epoxycarotenoid dioxygenase"
/protein_id="CAB10186.1"
/db_xref="GI:2769642"
/db_xref="SPTREMBL:O24023"
/translation="MATTTSHTATNTWIKTSLMPSKSEFGFASNSISLLKNQHNQSL
NINSSLAQPPILHFPKQSSNYQTPKNNTISHPKQNNSSSSSTSKWNLVQKAAAMAL
```



```
RESULT 8
LES439079      19018 bp      DNA      linear      PLN 13-MAR-2003
LOCUS          Lycopersicon esculentum nced1 gene for 9-cis-epoxycarotenoid
DEFINITION     dioxygenase.
ACCESSION      AJ439079
VERSION        AJ439079.2 GI:28974076
KEYWORDS       9-cis-epoxycarotenoid dioxygenase; nced1 gene.
SOURCE         Lycopersicon esculentum (tomato)
ORGANISM       Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 Thompson,A.J.
REFERENCE      Complementatation of notabilis, a tomato mutant deficient in the
AUTHORS        abscisic acid biosynthetic enzyme 9-cis-epoxycarotenoid dioxygenase
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 19018)
AUTHORS        Thompson,A.J.
TITLE          Direct Submission
JOURNAL        Submitted (13-MAR-2002) Thompson A.J., Plant Genetics and
                Biotechnology, Horticulture Research International, Wellesbourne,
                Warwick, CV35 9BF, UNITED KINGDOM
REMARK         revised by author [14-MAY-2002]
COMMENT        Related mRNA entry Z97215.
FEATURES       Location/Qualifiers
                1..19018
                /organism="Lycopersicon esculentum"
                /mol_type="genomic DNA"
                /cultivar="Money-maker"
                /db_xref="taxon:4081"
                /chromosome="7"
                /map="7q40"
                /clone="NCEd1.1"
                /rpt_family="tomato anionic peroxidase inverted repeat
                (TAPTR)"
repeat_region  267..567
                /rpt_type=INVERTED
                2343..3038
                /rpt_type=DIRECT
gene           5304..7570
                /gene="nced1"
TATA_signal   5304..5311
                /gene="nced1"
TATA_signal   5334..5337
                /gene="nced1"
mRNA          5404..7570
                /gene="nced1"
CDS           5482..7299
                /gene="nced1"
                /function="abscisic acid biosynthesis"
                /codon_start=1
                /product="9-cis-epoxycarotenoid dioxygenase"
                /protein_id="CAD30202.1"
                /db_xref="GI:28974077"
                /translation="MATTTSHATNTWTIKLSMPSSKEFGFASNSISLLKNQHNQRSL
                NINSLSQAPLILHPFKQSSNYQTTPKNNITISHPKQENNNSSSTSKWNLVOKAAMAL
                DAVESALT KHELEHPDKTADPRVQISGNFAPVPENPVCSLPVTGKIPKCVQGVYR
                NGANPLPEPTAGHFFDGDGMVHAVQFKNGSAYACRFTETERLVQEKALGRVPFPA
                IGEUHGSGIARLMLFYARGLFGLVDSKGTGVANAGLVFNNELLAMSEDDLPLPHVK
                VTPGDLKTEGRFDFDGLKSTMIATHPKLDPVSGELFALSVDVTKPLKXVFRFSKNG
                EKSNDVEIPVEDPTMMHDFALTENFVPIPDQGVFKMSEIRGSPVYDKNKVRSEFG
                ILDKYKDGSLKWEVDFDCFCFLHWNAAEEADEIIVIGSCMTPDPSIPNECDGL
                KSVLSRLNLTKGKTRKSIENPDEQVLEAGVARNKILGRTEYALAIAPFPK
                VSGFKNVLFTEGVEKFTYGDKNYGGEPFLFPRDPNSKEEDDGYILAFVHDEKEWKSE
                LQIVNANSKLKLEATVKLPSRVPYGFHGTFFINANDLANQA"
polyA_site    7570
                /gene="nced1"
                /evidence=experimental
repeat_region 8285..8575
                /rpt_family="tomato anionic peroxidase inverted repeat
                (TAPTR)"
```

```

QY 249 AlaGlyLeuValTyrPheAsnGlyArgLeuLeuAlaMetSerGluAspLeuProTyr 268
Db |||||
Db 6241 GCCGGTTAGTCTATTTCATTAACCGATTACTTGTCTGTAAGATGATTTGCGCTTAC 6300
QY 269 GlnValGlnIleThrProAsnGlyAspLeuLeuThrValGlyArgPheAspGly 288
Db |||||
Db 6301 CATGTAAGGTAAACACCCCGCGGATCTTAAACAGAGGGTGGATTCGATTTCCAGCGC 6360
QY 289 GlnLeuGluSerThrMetIleAlaHisProLysValAspProGluSerGlyGluLeuPhe 308
Db |||||
Db 6361 CAGCTAAATCCACCATGATAGCTACCCAAAGCTCGACCCAGTTTCGGGTGAGCTATT 6420
QY 309 AlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerProAsp 328
Db |||||
Db 6421 GCTCTTAGCTAGCATGCTGATTCAGAAAGCCATACCTCAAGTACTTCAGATTTTCAAAAAAT 6480
QY 329 GlyThrLysSerProAspValGluIleGlnLeuAspGlnProThrMetMetHisAspPhe 348
Db |||||
Db 6481 GGGGAAAATCAATGATGTTGAAATTCGATTTGAAGACCCCAACAAATGATGATGATTTTC 6540
QY 349 AlaIleThrGluAsnPheValValProAspGlnGlnValValPheLysLeuProGlu 368
Db |||||
Db 6541 GCAATTACTGAGAACTTCGTCGTCATTCTGTATCAACAGTCTGTTTCAAGATGCTGAA 6600
QY 369 MetIleArgGlyGlySerProValValTyrAspLysAsnLysValAlaArgPheGlyIle 388
Db |||||
Db 6601 ATGATCCGTCGAGGTTCACCGGTGTTTACGACAAAGACAAAGTTTCCCGATTGCGTATT 6660
QY 389 LeuAspLysTyrAlaGluAspSerSerAsnIleLysTyrIleAspAlaProAspCysPhe 408
Db |||||
Db 6661 CTGGATAAGTACGCCAAGATGGGTCTGATTTGAAATGGTGTGAAGTACCTGATTTGTTTC 6720
QY 409 CysPheHisLeuTrpAsnAlaTrpGluGluProGluThrAspGluValValIleGly 428
Db |||||
Db 6721 TGTTTCCACCTCTGGAATGCTTGGGAAGACGAGAAACAGATGAATCGTTGTAATGGT 6780
QY 429 SerCysMetThrProProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerVal 448
Db |||||
Db 6781 TCATGTATGACACCAACAGATCCATTTTCAATGAATGTGATGAAGGCTAAAGAGTGT 6840
QY 449 LeuSerGluIleArgLeuAsnLeuLysThrGlyGluSerThrArgProIleIleSer 468
Db |||||
Db 6841 TTATCCGAAATCCGCTCAATTTGAAACAGCGGAAATCAACAGAAATTCATAATCGAA 6900
QY 469 AsnGluAspGlnGlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArg 488
Db |||||
Db 6901 AACCCGGATGAACAGTGAATTTAGAGCTGGAATGGTGAACCGGAAACAAATCGGAAGG 6960
QY 489 LysThrLysPheAlaTyrIleuAlaLeuAlaGluProTrpProLysValSerGlyPheAla 508
Db |||||
Db 6961 AAAACAGAGTATGCTTATTGGCTATCGTGAACCATGCCCAAAAGTTTCTGTTTTCGA 7020
QY 509 LysValAspLeuThrThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGly 528
Db |||||
Db 7021 AAAGTAAACCTGTTCCCGGTGAAGTTGAGAAATTCATTTATGGTGAACAAATATGTT 7080
QY 529 GlyGluProLeuPheLeuProGlyGlu-----GlyGlyGluGluAspGluGlyTyrIle 546
Db |||||
Db 7081 GGGGAACCTCTTTTTTACCAAGACGCCCAACACAGGAGGAGAGACGATGGTTATATT 7140
QY 547 LeuCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaVal 566
Db |||||
Db 7141 TTAGCTTCGTTCCAGCATGAGAAAGATGGAATCAGAAATCAGAACTGCAAAATGTTTAAACGCAATG 7200
QY 567 SerLeuGluValGluAlaThrValLysLeuProSerArgValProTyrGlyPheHisGly 586
Db |||||
Db 7201 AGTTTGAAGTTGGAGCGCACTGTGAAGCTTCCATCAAGAGTTTCTTATGGATTTTCATGGA 7260
QY 587 ThrPheIleGlyAlaAspAspLeuAlaLysGln 597
Db |||||
Db 7261 ACATTCATAAACGCCAATGATTTGGCAATCAG 7293

```

# RESULT 9

## AX148316

### LOCUS

#### DEFINITION

#### AX148316

#### VERSION

#### AX148316.1

#### KEYWORDS

#### SOURCE

#### ORGANISM

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### FEATURES

#### source

#### CDS

#### Location/Qualifiers

#### 1..1818

#### /organism="Lycopersicon esculentum"

#### /mol\_type="genomic DNA"

#### /db\_xref="taxon:4081"

#### 1..1818

#### /note="unnamed protein product"

#### /codon\_start=1

#### /protein\_id="CAC41199.1"

#### /db\_xref="GI:14347204"

#### /translation="MATTTSHTATNTWIKTLMSPSSKEFFASNSISLKNHNRSL

#### NINSSLOAPPIHFHKOSSNYQTPKNNTISHPKOENNNSSSSSSKNWLQKAAAMAL

#### DAVSATLKELEHPLPKTADPRVQISGNFAPVPENPVCQSLPTVKIKPKVQGVYVR

#### NGANLPEPTAGHHFFDGDGNVHAVQFKNGSASACRFTETERLVQEKALGRVPVPPKA

#### IGELHSGSGLARMLFARGLVGHDHSGTGVANAGLVYFNRLLLAMSEDLPLPHVK

#### VTPTGDLKTEGRFDQGLKSTMLAHPKLDPSVGLFALSVDVIQPKYLKFRFSKNG

#### EKSNVDIPEVEDPTMHDFAITENFVVIPOQVVFVKMSEMRGGSVVVDKMKVSRFG

#### ILDKYAKGSDLKWVEVDCFCFLHNAWEAEDELIVIGSCMTPPDSINECDEGL

#### KSVLSRLNLKTKSTRKSIENPDQVNLEAGMNRNKLGRKTEYAYLALAEWPVK

#### VSGFAKNLFTGEVEKFIYGDNKYGGEPFLFPRDPNRSKBEDDDGYILAEVHDEKEWKSE

#### LQIVNMSLKLEATVKLPSRVPYGFHGTFINANDLANQA"

#### BASE COUNT

#### 569 a 383 c 369 g 497 t

#### ORIGIN

#### Alignment Scores:

#### Pred. No.:

#### Score:

#### Percent Similarity:

#### Best Local Similarity:

#### Query Match:

#### DB:

#### US-09-758-269-6 (1-599) x AX148316 (1-1818)

#### QY 5 ThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThrGlnProLeu 24

#### Db 7 ACTACTACTTCACATGCCACCAATACATGGAAT-----AAGACTAAGTTG 51

#### QY 25 SerSerGlnSerSerAspLeuSerTyrCysSer-----SerLeuProMetAla 41

#### Db 52 TCAATGCCATCATCAAGGAGTTGGTTTTCATCAAACTCTATTCTCTACTCAAAAT 111

#### QY 42 SerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProProAlaLeuHis 61

#### Db 112 CAACATAATAGGCAAGTCTCAACATTAATTCCTCTCTTCAAGCTCCACCTATCTCAT 171

#### QY 62 PheProLysGlnSerSerAsn-----SerProAla-----IleValValLysProLys 77

#### Db 172 TTTCTTAACCAATCTTCAATATCAACACCAAGAAATAATACATTTTCACACCAAAA 231

#### QY 78 AlaLysGluSerAsn-----ThrLysGlnMetAsnLeuPheGlnArgAla 92

#### Db 232 CAAGAAACAACAACACTCTCTCTTCTTCACTTCCAAAGTGAATTTAGTCGAGAAGCA 291

#### QY 93 AlaAlaAlaLeuLeuAspAlaAlaGlyPheLeuValSerHisGluLysLeuHisPro 112



```
Db 292 GCAGCAATGGCTTTAGATGCTGTAGAAAGTGCTTTAACTAAACATGAACCTTGAACACCCCT 351
Qy 113 LeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValAsnGlu 132
Db 352 TTGCCGAAACAGCCGACCCAGAGTCAGAGTTCTCGGAAATTTTCTCGGTACCGGAA 411
Qy 133 GlnProValArgAsnLeuProValValGlyLysLeuProAspSerIleLysGlyVal 152
Db 412 AATCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 471
Qy 153 TyrValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHisHisPhePheAsp 172
Db 472 TACGTTCCGAAACGAGGCTAACCTCTTTTGAACCAACCCGCGACACCACTTTCTTCGAC 531
Qy 173 GlyAspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyrAlaCysArg 192
Db 532 GCGCAGCGTATGGTTACGCGCTCAATTCACCAAAATGGTCCGCTAGTTACGCTTGGCGT 591
Qy 193 PheThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPheProLys 212
Db 592 TTCACTGAACAGAGAGCTTTGTTCAAGAAAAGCTTTGGGTCCGCTCTTTTCCCTAAA 651
Qy 213 AlaIleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArg 232
Db 652 GCCATTGGTGAATTACATGGTCACTCTGGAATTGCAAGGCTTATGCTGTTTACGCTCGT 711
Qy 233 AlaAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuVal 252
Db 712 GGGCTCTTCGGACTTGTGTATGATCAGTAAGAAAGAACTGGTGTGCAAAACGCCGGTTAGTC 771
Qy 253 TyrPheAsnGlyArgLeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIle 272
Db 772 TATTTTCAATAACCAATACCTGCTATGCTGCTGAAAGATGATTTTGGCTTACCATGTAAAGGTA 831
Qy 273 ThrProAsnGlyAspLeuLysThrValGlyArgPheAspPheAspGlyGlnLeuGluSer 292
Db 832 ACACCCCGCGGATCTTTAAACAGAGGGTTCGATTTCGATTCGACGGCCAGCTAAATCC 891
Qy 293 ThrMetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeuSerTyr 312
Db 892 ACCATGATAGTACACCAAGCTCGACCACTTTCCGGTGAGCTATTTCGCTTCTTAGCTAC 951
Qy 313 AspValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSer 332
Db 952 GATGTGATTCAGAGCCATACCTCAAGTACTTCAGATTTTCAAAATATGGGAAAAATCA 1011
Qy 333 ProAspValGluIleGlnLeuAspGlnProThrMetHisAspPheAlaIleThrGlu 352
Db 1012 AATGATGTTGAAATTCAGATTGAAGACCAACAATGATGATGATTTTCGCAATTTACTGAG 1071
Qy 353 AsnPheValValValProAspGlnGlnValValPheLysLeuProGluMetIleArgGly 372
Db 1072 AACTTCGTGCTGCTATTCCTGATCAACAGTCGTTTTTCAAGATGTCTGAAATGATCCGTGGA 1131
Qy 373 GlySerProValValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyr 392
Db 1132 GGTTCACCGGTGGTTTACGACAGAACAAAGTTTCCCGATTGGTATTCTGGATAAGTAC 1191
Qy 393 AlaGluAspSerSerAsnIleYrTrpIleAspAlaProAspCysPheCysPheHisLeu 412
Db 1192 GCGAAAGATGGTCTGATTTGAAATGGGTTGAAGTACCTGATGTTTCTGTTTCCACCTC 1251
Qy 413 TrpAsnAlaTrpGluProGluThrAspGluValValIleGlySerCysMetThr 432
Db 1252 TGGAAATGCTTGGGAGAGACGAAACAGATGAAATCGTTGTAATTTGGTTCATGTATGACA 1311
Qy 433 ProProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIle 452
Db 1312 CCACAGACTCCATTTTCAATGAATGTGATGAAGGGCTTAAAGAGTGTTTTATCCGAATC 1371
Qy 453 ArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGln 472
Db 1372 CGTCTCAATTTGAAAAACAGGGAATCAACAAGAAAAATCCATAATCGAAAAACCCCGGATGAA 1431
```

```
Qy 473 GlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPhe 492
Db 1432 CAAGTGAATTTAGAACTGGAAATGGTGAACCGAAACAAACTTCGGAAGAAACAGAGTAT 1491
Qy 493 AlaTyrLeuAlaLeuAlaGluProTyrProLysValSerGlyPheAlaLysValAspLeu 512
Db 1492 GCTTATTTGGCTATCGCTGCAACCATGGCCAAAGATTCTCGTTTTGCAAAAGTAAACCTG 1551
Qy 513 ThrThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGluProLeu 532
Db 1552 TTCACCGGTGAAGTTGAGAAATTCATTATATGGTGACAAACAATATATGTTGGGAACCTCTT 1611
Qy 533 PheLeuProGlyGlu-----GlyGlyGluGluAspGluGlyTyrIleLeuCysPheVal 550
Db 1612 TTTTACCAAGAGACCCCAACAGCAGAGAAAGACAGCATGCTATATATTTTATGCTTTCGTT 1671
Qy 551 HisAspGluLysThrTyrLysSerGluLeuGlnIleValAsnAlaValSerLeuGluVal 570
Db 1672 CACGATGAGAAAGATGGAATCAGAACTGCAAAATTTGTTAACGCAATGAGTTTGAAGTTG 1731
Qy 571 GluAlaThrValLysLeuProSerArgValProTyrGlyPheHisGlyThrPheIleGly 590
Db 1732 GAGCAACTGTGAAGCTTCATCAAGAGTTCCCTATGCAITTCATGGAACATTATTAAC 1791
Qy 591 AlaAspAspLeuAlaLysGln 597
Db 1792 GCCAATGATTGGCAATCAG 1812

RESULT 10
LOCUS BD017436 1818 bp DNA linear PAT 27-AUG-2002
DEFINITION Transgenic plant using neozanthine cleaving enzyme gene.
ACCESSION BD017436
VERSION BD017436.1 GI:22558612
KEYWORDS Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
REFERENCE 1 (bases 1 to 1818)
AUTHORS Iuchi, K., Kobayashi, M. and Shinozaki, K.
TITLE Transgenic plant using neozanthine cleaving enzyme gene
JOURNAL Patent: JP 2001258579-A 8 25-SEP-2001;
THE INSTITUTE OF PHYSICAL AND CHEMICAL RESEARCH
COMMENT OS Lycopersicon esculentum (tomato)
PN JP 2001258579-A/8
PD 25-SEP-2001
PF 11-JAN-2001 JP 2001003476
PI KIYOSHI IUCHI, MASATOMO KOBAYASHI, KAZUO SHINOZAKI PC
C12N15/09, A01H5/00, C12N5/10, C12N5/02, C12N15/00, C12N5/00 CC
Transgenic plant using neozanthine cleaving enzyme gene FH Key
Location/Qualifiers
(1) . (1818).
FEATURES
source
Location/Qualifiers
1. .1818
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/db_xref="taxon:4081"
BASE COUNT 569 a 383 c 369 g 497 t
ORIGIN
Alignment Scores:
Pred. No.: 1,24e-156 Length: 1818
Score: 2280.50 Matches: 435
Percent Similarity: 82.04% Conservative: 63
Best Local Similarity: 71.66% Mismatches: 90
Query Match: 72.40% Indels: 19
DB: 6
US-09-758-269-6 (1-599) x BD017436 (1-1818)
```



Qy	5	ThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThrGlnProLeu	24
Db	7	ACTACTACTTCACATGCACAAATACATGGATT-----AAGACTAAGTTG	51
Qy	25	SerSerSerGlnSerSerAspLeuSerTyrCysSer-----SerLeuProMetAla	41
Db	52	TCAATGCCATCATCAAGAGAGTTGGTTTGTGATCAAACTCTATTTCCTACTCAAAAT	111
Qy	42	SerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProProAlaLeuHis	61
Db	112	CAACATAATAGCAAGTCATCAACATTAATTCCTCTCTTCAAGCTCCACCTATATCTTCAT	171
Qy	62	PheProLysGlnSerSerAsn-----SerProAla-----IleValValLysProLys	77
Db	172	TTTCTTAAACAATCTTCAAATTTATCAAAACACCAAGAAATAATACAAATTTCACACCAAAA	231
Qy	78	AlaLysGluSerAsn-----ThryLysGlnMetAsnLeuPheGlnArgAla	92
Db	232	CAAGAAAAACAACACTCCTCTCTTCCTCAACTTCCAAGTGGAAATTTAGTCAGAAAGCA	291
Qy	93	AlaAlaAlaLeuAspAlaAlaGluGlyPheLeuValSerHisGluLysLeuHisPro	112
Db	292	GCAGCAATGGCTTTAGATGCTGTAGAAAGTGCTTTAACTAAACATGAACCTTGAACACCCCT	351
Qy	113	LeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValAsnGlu	132
Db	352	TTGCGAAAAACAGCCGACCCACGAGTCCAGATTTCTGGGAATTTTGTCTCCGTACCGGAA	411
Qy	133	GlnProValArgArgAsnLeuProValValGlyLysLeuProAspSerIleLysGlyVal	152
Db	412	AATCCAGTCTGTCAATCTCTCCGGTCAACCGGAAAAATACCCAATGTGTTCAGAGCGCT	471
Qy	153	TyrValArgAsnGlyValAsnProLeuHisGluProValThrGlyHisHisPhePheAsp	172
Db	472	TACGTTCCAAACGGAGCTAACCCCTCTTTTGAAACCAACCGCGGACACCATTTCTTCGAC	531
Qy	173	GlyAspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyrAlaCysArg	192
Db	532	GGCAGCGGTATGGTTTACCGCGCTTCAATTCAAAAATGGTTCGGCTAGTTACGTTGCCGT	591
Qy	193	PheThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPheProLys	212
Db	592	TTCTACTGAAAACAGAGAGCTTGTGTTCAAGAAAAAGCTTTGGGTGCGCCCTGTTTCCCTAAA	651
Qy	213	AlaIleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArg	232
Db	652	GCCATTGGTGAATTCATGGTCACTCTCGAATTCGAAGGCTTATGCTGTTTTAGCTCGT	711
Qy	233	AlaAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAlaAsnAlaGlyLeuVal	252
Db	712	GGGCTCTTCGGACTTGTGTGATCAACAGTAAGGAACCTGGTGTTCGAAACGCCGGTTTAGTC	771
Qy	253	TyrPheAsnGlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIle	272
Db	772	TATTTCAATAACCCGATTACTTGGCTATGTCTCAAGATGATTTGCGCTTACCATGTAAGGTA	831
Qy	273	ThrProAsnGlyAspLeuLysThrValGlyArgPheAspPheAspGlyGlnLeuGlnSer	292
Db	832	ACACCCACCGCGATCTTAAACACAGAGGTCATTCGATTTTCGACGCCGACGTAATATCC	891
Qy	293	ThrMetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeuSerTyr	312
Db	892	ACCATGATAGCTCACCCAAAGCTCGACCCAGTTTCCGGTGAGCTATTTTGCTCTTAGCTAC	951
Qy	313	AspValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSer	332
Db	952	GATGTGATTCAGAGCCATCACTCAAGTACATTCAGATTTTCAAAAAATGGGGAAAAATCA	1011
Qy	333	ProAspValGluIleGlnLeuAspGlnProThrMetMetHisAspPheAlaIleThrGlu	352
Db	1012	AATGATGTTGAAATTCNGTTGAAGACCCCAACATGATGCAATGATTTTCGCAATTACTGAG	1071
Qy	353	AsnPheValValProAspGlnGlnValValPheLysLeuProGluMetIleArgGly	372



```

Db      1597 GTGAACCTTTGAAGCCGCGCATGTTGAACAGAAACAAGCTAGGAAAGGAAACACTCAGTTCGCT 1656
Qy      494 TyrLeuAlaLeuAlaGluProTyrProLysValSerGlyPheAlaLysValAspLeuThr 513
Db      1657 TATCTGGCTCTCGCGAGCCCTGGCCCAAGTCTCGGGCTTTGCCGAAGTGTATTGTTTC 1716
Qy      514 ThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGluProLeuPhe 533
Db      1717 AGTGGGGAAGTGCAGAAAGTACATGATGAGAGAAGAAAGATTCGGTGGGAGCCTCTGTTT 1776
Qy      534 LeuProGlyGluGlyGlyGluGluAspGluGlyTyrLysLeuCysPheValHisAspGlu 553
Db      1777 CTTCCTCC---AACGGCGAAGAAAGGTGATGGGTATATTCTGGCATTCTGTCACACGAG 1833
Qy      554 LysThrTrpLysSerGluLeuGlnIleValAsnAlaValSerLeuGluValGluAlaThr 573
Db      1834 AAGAGTGGAAATCCGAGCTGCAGATTGTGAATGCCCAAAATTTAAAGCTCGAAGCTTCC 1893
Qy      574 ValLysLeuProSerArgValProTyrGlyPheHisGlyThrPheIleGlyAlaAspAsp 593
Db      1894 ATCAAACTCCCTCTCGCGTCTTCATGGTTTTCATGGCAGCTTTTATTCAATCTAAGGAT 1953
Qy      594 LeuAlaLysGln 597
Db      1954 TTGAGGAAGCAA 1965

RESULT 12
LOCUS   AX652128 95769 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 1979 from Patent WO03000898.
ACCESSION AX652128
VERSION   AX652128.1 GI:29154942
KEYWORDS
SOURCE   Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Karagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, F. and Zou, G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 03000898-A 1979 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES
source
1. 95769
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT 30958 a 16961 c 16682 g 31168 t
ORIGIN
Alignment Scores:
Pred. No.: 3 456-146 Length: 95769
Score: 2170.50 Matches: 419
Percent Similarity: 80.61% Conservative: 80
Best Local Similarity: 67.69% Mismatches: 87
Query Match: 68.90% Indels: 33
DB: 6 Gaps: 9

US-09-758-269-6 (1-599) x AX652128 (1-95769)
Qy 1 MetAlaSerPheThrAlaThrAlaAlaValSerGlyArgTrpLeu----- 15
Db 33716 ATGGCTTCTACTACCTTACTCTCTTCTACTTCAACCCCAATTCCTCGACCGAATCTTCTCA 33775
Qy 16 ---GlyGlyAsnHisThrGlnProProLeuSerSerSerGlnSerAspLeuSerTyr 34
Db 33776 ACTTCATCTCTCTCTCGAGACCTAACTTCAATCT-----CTATCTTTT 33820
Qy 35 CysSerSerLeuPro-----MetAlaSerArgValThr 45

```

```

Db      33821 TCCTCAACACCTTAGAAGAACAGAAACTAGTAGTCCCTCTGATGTTCTTCTCCTCCGTCAAC 33880
Qy      46 ArgLysLeuAsnValSerSerAlaLeuHisThr-----ProProAlaLeuHis 61
Db      33881 AAAAAATCTTCGGTCTCGTCTCTCTCAATCTCCGACATTCAAACACCGCTCTGG--- 33937
Qy      62 PheProLysGlnSerSerAsnSerProAlaIleValLysProLysAlaLysGluSer 81
Db      33938 -----AAGAAACTATGTAAACGATGTCAACCACTTGATC---CCCAAAACGACGAATCAG 33988
Qy      82 AsnThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaGlu 101
Db      33989 AATCAAAA-----TAAACCCGCTGCAGAGAACCCGCGCTATGGTTTTAGACGCGGTGAA 34045
Qy      102 GlyPheLeuValSerHisGlu---LysLeuHisProLeuProLysThrAlaAspProSer 120
Db      34046 AACGCCATGATCTACACGAGCGCGCGCTCATCCCATCTCTAAACCGCGGATCCTGCC 34105
Qy      121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgAsnLeuPro 140
Db      34106 GTCCAAATAGCGGGAACCTTCTCCGCTACCGGAGAAACCTGCTCGTCAATCAACCTCG 34165
Qy      141 ValValGlyLysLeuProAspSerLysGlyValTyrValArgAsnGlyAlaAsnPro 160
Db      34166 GTGACTGGAAACAGTACACAGAAATGCATTCAAGGAGTTTACGTCAAGAAACGGAGCAATCCA 34225
Qy      161 LeuHisGluProValThrGlyHisPheAspGlyAspGlyMetValHisAlaVal 180
Db      34226 CTTCAAAACAGTCTCCGGCCACCATTTGTTCCAGCGGAGCGGTATGGTACACGCTGTC 34285
Qy      181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200
Db      34286 CGGTTTCGATAACGCGTCTGTTAGTACGCTTGTGCGTTTACCGAAACAAACCGGTGTTT 34345
Qy      201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220
Db      34346 CAAGAACGTGAATGTGTCGTCGCGTTTCCCAAAAGCAATTTGGAGAGCTTCACGAGCAT 34405
Qy      221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspPro 240
Db      34406 TTAGAAATCGCAAGCTTATGCTCTTCAATACCCGCGGGCTATTGGGTATTGTCAGCCCG 34465
Qy      241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuAla 260
Db      34466 ACCGAGGAGCTCGGTGTCGCTAACCGCGTTTATTATTTCAATGGTCACTCTCTAGCC 34525
Qy      261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280
Db      34526 ATGTCGGAAGACGATTTACCGTACCATGTCAAAAGTTACTCAAAACCGGAGATTTAGAAACT 34585
Qy      281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300
Db      34586 TCGSGTTCGTACGATTTTCGACGCTCAGTTAAATCAACATGATAGCCACCCCGGAAATC 34645
Qy      301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeu 320
Db      34646 GATCCGAAACCCCGAGAACTCTTCGCTTAAAGTACGACGCTCGTTTCAAGACCTTACTTA 34705
Qy      321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340
Db      34706 AAATACTTCAGATTCAATCCGACGCGTGAGAAATCACACAGACGTGAGATTCGCGTTGAT 34765
Qy      341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValValProAspGln 360
Db      34766 CAACCAACGATGATCCAGGATTTCCGATCTACTGAGAAATTTCTGATGATTTCCAGACCAA 34825
Qy      361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValValTyrAspLys 380
Db      34826 CAAGTGGTATTTCGATTCGCGAGATGATAAGAGGTGGTTCTCCGCTGGTTTACGACGAG 34885
Qy      381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnLys 400
Db      34886 AAGAAAGAAATCAAGATTTCGGGATTTGAATAAAAACCGGAAAGATGCTTCGTTCGATTCAA 34945

```



CDS

```
join(8524..8759,8846..9287)
/notes="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F3f9.4"
/protein_id="AAF71794.1"
/db_xref="GI:8052530"
/translation="MEALASIFKLSLLTLLSGGSLDLSLPPNCNRAECPEYEVVHA
GNGYEIHRVTWTTWISTEPIQDISNEASGNGNQLSDYMGNNYDHYORIEALPYIT
QYQNSLSTIVSFVFPKAFQPPPGNNLHVQRWDSRYVAVKQISGYVADHRIGKQVA
ELKASLQGTWMAKAEKSRGTGSGAWAYTAQFSWPQWSQRVNEIWFPEMEDEE
TVSIVTQ"
join(13444..13814,13956..14310,14399..14662)
/notes="similar to dioxygenase protein"
/codon_start=1
/evidence=not_experimental
/product="F3f9.5"
/protein_id="AAF71795.1"
/db_xref="GI:8052531"
/translation="WAVLSKPVAIKPSGSLIPVIDMSPESHALVKACEDFGPFVKV
INHGSAELVLEHETVDFSLPKSEKTVAGYFPGYGNKIGRNGDVGVWEYLLMN
AHDGSGFLPFLSLKSGTFRNALSEYTSVRKMTFDVLEKIDGLGKPRNTLSKL
VSDQNTSLRLNHYPPCLSNKKTNGKKNVIGFGEHTDPOIISVLRSNNTSGLOINL
NDGSTSVPDDHTSFEFNVGDSIQWMTGRFKSVPRHVLNCKKSRVSIYPAGPSLT
QRIAPLTCLIDNEDRLYEFTWSEYKNSYNSRUSDNRQQOFRKTIKNLNL"
complement(join(15755..16715,16810..16829))
/notes="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F3f9.6"
/protein_id="AAF71812.1"
/db_xref="GI:8052548"
/translation="MPKRLGIRGSELPGQSRPLRTSLSTSSDPHLSRPIIDRSP
KLGDRRSPRSGPHTDPLSKGLSGLSQAOEELRLKQOLAKAEAAKRP
AOEELHKKSKPNTAPERDDIPGDGHOETDVFVLEDEKASEKTKNDLASELE
INVULARLYDLSEKRVSLSENETLKQLKKTDTWMSCKAKEDIASKVSQIGSELE
ESNETAKLKKLSEVEENETLEAEMLKLVQTEQWKAADAALVSGVGMNGRF
SQDCSMKHFAGRFVSGMADDDSGGKKSKGKNFGDLWKKQKQ"
complement(join(19162..19689,19767..19833,20044..20164,
20234..20378,20482..20623,20693..20738,20826..20877,
20965..21156))
/notes="unknown protein; similar to ESTs gb|T41672.1,
gb|A1992710.1, and gb|AAS86078.1"
/codon_start=1
/evidence=not_experimental
/product="F3f9.7"
/protein_id="AAF71811.1"
/db_xref="GI:8052547"
/translation="MGNLGRKQVVEERYTKPQGLYVNVKDVVKLARKLIVESKLAP
CYPGDESDHLEECPIFLYYPNLRNRCCKMSICTECFLQMNPNRARPQKFLIN
LMGFFPFFVSLTFMRPCFKTPNAYEYRGVSKSKEEGIEQVEORVIEAKIRMQKE
MODDEKMKRLSESSSSTAMTGMWEGYSGFLFVPMFAAISYNLSMDDEGIEAPSQ
NADVROHSRPNRGNREDDVDLELMWMEALWLSQETGORTNRSAGEITSSROYVT
DNHSYSSPPRVTPIVEPATPSSSGGLSCALSALAERQMVGESSSHNNHNVSSY
SNLPGNCSDYDIEQVDGIDNHHHRRHHYEMGETGSSNSYSSYMTGEGFNFPPPP
PLVIVPESFEQMMAMAVSMAEHATTICAPTEVTWQ"
complement(join(22808..23063,23118..23244,23511..23706))
/notes="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F3f9.8"
/protein_id="AAF71810.1"
/db_xref="GI:8052546"
/translation="MSGKGVKSEPMKVVFINTQVETDARSFTVQELTKNAIVA
AGPSPSPAFDRCDGDSKIGIPINVFWKVMISYLSVRPMSNPKRAYISFHMFLFK
TRMKQNTQNKKSCVTKALKATRLRSQASTMSGLWGVFATQVQLGTFPSIFFI
YLFWFPFSSCKMFLTKRILISLKLVSILKTV"
join(26752..26907,27020..27871,27923..28153)
/notes="similar to exopolysaccharuronase precursor"
sp|Q00293|PGLX_ASPTU"
/codon_start=1
/evidence=not_experimental
/product="F3f9.9"
```

CDS

CDS

```
/protein_id="AAF71796.1"
/db_xref="GI:8052532"
/translation="MVSTISGSLILFIAAVALSSISAAPSAALVGRKVDVRSYGARG
DKDNTNTWAKTDQWGLPRVYIPFGFYLGAVAFPGTCKSRISFIKGTLLA
PKDPNAIKODSMWIFRYDYLTGSGGLDGGGYSWPLNCRQTHNCRALPMNMGFO
FVRPSRLTRIKSNKMGHLAFPSVQHEIDITRVNIKAPGDSPTDGIKIGSSNHHKIH
HYDITGDDCTAILSGTFLNDINKNCGPHGHSVGSUGKFKGKESVQGLIVRNSIFN
GTSGNVRITKTPWSPGPNLVSFLNQLMIDVQSPINIDQRYCPNPPCSFQVTSLTR
KNSFSKIQIRDVKFQNIWGTSTAKEAVLQCSKNVPCKNVOLFNIHVHRGRDPATS
VCENVGGWIGGKISPPSCIR"
33566..35539
/notes="similar to 9-cis-epoxycarotenoid dioxygenase
gp|AAF26356.1"
/codon_start=1
/evidence=not_experimental
/product="F3f9.10"
/protein_id="AAF71797.1"
/db_xref="GI:8052533"
/translation="MTLITISGMYYISLLSODAHHSQVQONTNVLVKKPIPKPQTAA
FNOESTMASTTLTLPSTQFLDRFTSSSRSPKQLQSLFSSTLRNKKLVVPCYVSS
SVNKKSSVSSLSQSPSTFPKPPQWKKLNDVTNLI PKTTONKLPVQTAAMVLDAVE
NAMISHERRHPHPKPTAPVQIAGNFPVPEKPVVHLNPTVGTVECIQGVYVRNGA
NPLHKPVSGHLEFDGDMHVARPDNGSVSVACRTEINRLVQERECRPPVPEKAGE
LHGLGIKALMLFNTRGLFGLVDPGTGGLGVANAGLVYFNGLHLSAMSEDDLPHYKVTQ
```

CDS

```
Alignment Scores:
Pred. No.:      Length:      95769
Score:          2170.50      Matches:      419
Percent Similarity: 80.61%      Conservative: 80
Best local Similarity: 67.63%      Mismatches: 87
Query Match:      68.90%      Indels:      33
DB:               Gaps:      9

US-09-758-269-6 (1-599) x AC013430 (1-95769)

Qy      1 MetalaserPheThrAlaThraAlaValSerGlyArgTrpLeu----- 15
Db      33716 ATGGCTTCTACTACCTTACTTCTTCTTCAACCAATCTCTCGACGAACTTTCTCA 33775

Qy      16 --GlyGlyAsnHisThrGlnProLeuSerSerSerGlnSerSerAspLeuSerTyr 34
Db      33776 ACTTCATCTTCTCTTCGAGACCTAACTCAATCT-----CTATCTTTT 33820

Qy      35 CysSerSerLeuPro-----MetalaserArgValThr 45
Db      33821 TCCTCAACACTTAGAAACAAGAACTAGTAGTCCCTGCTGTATGTTTCTTCTCCGTCAC 33880

Qy      46 ArgLysLeuAsnValSerSerAlaLeuHisThr-----ProAlaLeuHis 61
Db      33881 AAAAATCTTCGTCGTCGTCCTCTCTCTCAATCTCGACATTCAAACCAACCGCTTGG--- 33937

Qy      62 PheProLysGlnSerSerAsnSerProAlaIleValValLeProLysAlaLysGluSer 81
Db      33938 -----AGAAACTATGTAAAGCATGTCAACCACTTGATC---CCCAACACGACGAATCAG 33989

Qy      82 AsnThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaGlu 101
Db      33989 AATCCAAAA---TTAAACCCGGTGCAGAGAACCCGGCGGTATGGTTTATAGCCGGTGGAA 34045

Qy      102 GlyPheLeuValSerHisGlu---LysLeuHisProLeuProLysThrAlaAspProSer 120
Db      34046 AACGCCATGATCTCACAGCGCGCGCTCATCCCATCTCTTAAACCGCGGATCCTGCC 34105

Qy      121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgAsnLeuPro 140
Db      34106 GTCCAAAATAGCCGGGAATCTTCCCGGTACCGGAGAACCTGTCTGTCATACCTTCGG 34165

Qy      141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160
Db      34166 GTGACTGGACAGTACAGATGTCATTCAAGGAGTTTACGTACAGAAACGAGCAATCCA 34225

Qy      161 LeuHisGluProValThrGlyHisHisPhePheAspGlyAspGlyMetValHisAlaVal 180
Db      161 LeuHisGluProValThrGlyHisHisPhePheAspGlyAspGlyMetValHisAlaVal 180
```

```
Db 34226 CTTCAAAACAGCTCCGGCCACCATTGTTGTCAGCGAGAGCGGTATGGTACACGCTGTC 34285
Qy 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200
Db 34286 CGGTCGATAAAGCGGTCGGTTAGCTACGCTTGTCCGGTTTACCGAAACAAACCGGTGGTT 34345
Qy 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220
Db 34346 CAAGACGTGAATGTGGTCGTCGGTTTCCCAAGCAAGCAATGGAGAGCTTCAGGCAT 34405
Qy 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaGlyIleValAspPro 240
Db 34406 TTAGGAATCGCAAGCTTATGCTCTTCAATACCGCGGGCTATTGGGTAGTCGACCG 34465
Qy 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260
Db 34466 ACCGAGGACTCGGTGTCCTAACCGCGTTAGTTATTTCAATGGFCATCTCTTAGCC 34525
Qy 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280
Db 34526 ATGTCGGAAGACGATTACCGTACCATGTCAAGTTACTCAACCGGAGATTTAGAACT 34585
Qy 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300
Db 34586 TCGGTGCGGTACGATTTCGACGGTCAGTTAAATCAACATGATAGCCACCCGAAATC 34645
Qy 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrIleu 320
Db 34646 GATCCGAACCCGAGAACTCTTCGCTTAAAGTACGACGTCGTTTCAAAACGCTTACTTA 34705
Qy 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340
Db 34706 AATACTTCAGATTACATCCGCGGTGAGAAATCACCAGACGTCGAGATTCCGCTGAT 34765
Qy 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGln 360
Db 34766 CAACCAACGATGATCCAGATTTCGCGATCACTGAGAAATTCGTAGTGTATCCAGACCA 34825
Qy 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380
Db 34826 CAAGTGGTATTCCGATTCCCGGAGATGATAAGAGTGGTTCTCCGGTGGTTTACGACGAG 34885
Qy 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400
Db 34886 AAGAAGAAATCAAGATTCGGATTGTTGAATAAACCGCAAGATGCTCTCGCATTCAA 34945
Qy 401 TrpIleAspAlaProAspCysPheHisLeuTrpAsnAlaTrpGluGluProGlu 420
Db 34946 TGGATCGAAGTACCAGATTGTTTGTGTTTCCATCTATGGAATCTCTTGGGAAGAACCA 35005
Qy 421 ThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440
Db 35006 ACACACGAGGTTGTCGTGATTGGATCATGATGATGATGATGATGATGATGATGATGAT 35065
Qy 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460
Db 35066 CACGACGAACACATTCAGAGTGTGTTGTCGAGATAGCTTAAACCTTAAACAGGGAA 35125
Qy 461 SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480
Db 35126 TCAACACGTAGACCGGTTATCTCG-----GAACAAGTTAATCTCGAAGCTGGTATG 35176
Qy 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500
Db 35177 GTCAACCGGAATTTATTAGTAGTAAAAACCCGGTATGCTTATCTGGCTTTTAAACCGAAC 35236
Qy 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis 520
Db 35237 TGGCTTAAGTGTCCGGTTTCCCTAAGTGACATTATCGACCGGAGATTCGAAAGTAT 35296
Qy 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGlyGlu 540
Db 35297 ATTTACGAGAGAGGGGAAATACGAGGAGAGCTCTGTTTCTACCTTCGCGTGACGAGAA 35356
```

```
Qy 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560
Db 35357 GAAGACGAGGTTACATAATAGTGTGTTGTTTACACGAGGAGAGAGGTGAATTCGAACTC 35416
Qy 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580
Db 35417 CAACTCATAAACCGCTTAATATGAAGCTTGAAGCTACCGTAACGCTTCGTCGAGAGTG 35476
Qy 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGlnValVal 599
Db 35477 CTTTACGTTTCCACGAGAGCTTCATCAGTAAGAGAGATTATCGAAGCAAGCTTTG 35533

RESULT 14
AX148312
LOCUS AX148312 1839 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 11 from Patent EP1116794.
ACCESSION AX148312
VERSION AX148312.1 GI:14347199
KEYWORDS Vigna unguiculata (cowpea)
SOURCE Vigna unguiculata
ORGANISM Vigna unguiculata
REFERENCE 1
AUTHORS Iuchi, S., Kobayashi, M. and Shinozaki, K.
TITLE Transgenic plants carrying neoxanthin cleavage enzyme gene
JOURNAL Patent: EP 1116794-A 11 18-JUL-2001;
Riken (JP)
FEATURES
source 1..1839
Location/Qualifiers
/organism="Vigna unguiculata"
/mol_type="genomic DNA"
/db_xref="taxon:3917"
1..1839
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC41197.1"
/db_xref="GI:14347200"
/translators="MPSSASNTWFNATLPSPPFKDLPSSTNLLPLRKTSSNTIT
CSLQTLHPKQYQSTSTSTATTTPKTKTTTTPRETNPISDTPQIPQKW
NFIQKAATALDVELALVSHRKHPLPKTADPRVOIAGNAPVPEHAADQGLPVQK
IPKCIDGVYRNGANPLYPVAGHHFDGGMVHAKFTNGAASVACRFTETORLSOE
KSUGRPVFPKAIAGELHGHSGIARLLFYARGLFGLVDGSGQGMVANGLVFNHLLA
MSDDDLUPYHRIITPNDGLTTRVDFNGQLNSTMIAHPKLDIPVDGLHALSYDVIQK
YLKYFRSPDGKSPDVEIPLKEPTWMHDFAITENFVVPDQVVFKLTEMITGGSVP
VYDKNTRRREGLHKNAKANDANRWIDAPDFCFHLMWNAWEEPEEVEVIGSMTQA
DSIFNECEESLSVLSIRLNLTGKSTRPIISDAQVNLEAGMVRNKLGRKTQA
YLAALAEPPKVSQFAKVDLLSGVKMYGEEKGEGPLPLPNGKEDDGLVILAFVHD
EKWAKSELQVNAQNLKLEASIKLPSPVPGFHGTFHISKDLRKQA"
BASE COUNT 447 a 595 c 442 g 355 t
ORIGIN

Alignment Scores: 1.78e-148 Length: 1839
Pred. No.: 2168.50 Matches: 423
Score: 78.44% Conservative: 61
Percent Similarity: 68.56% Mismatches: 98
Best Local Similarity: 68.84% Indels: 35
Query Match: 6 Gaps: 8
DB:

US-09-758-269-6 (1-599) x AX148312 (1-1839)
Qy 8 AlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThrGlnProProLeuSerSerSer 27
Db 7 TCATCAGCTTCAACACTTGGTTTAAACGCCACACTCCCATCTCCCTTCAAGACCTA 66
Qy 28 GlnSerSerAspLeuSerTyrCysSerSerLeuProMetAlaSerArgValThrArgLys 47
Db 67 CTTTCCACATCT---TCTCCCAACAACTTACTTCTCTTTA-----AGGAAACATCTCT 117
```

QY 48 LeuAsn---ValSerSerAlaLeuHisThrProProAlaLeuHisPheProLysGln--- 65  
DB 118 TCCAACACCATCAGTGTTCCTTCAACA-----CTCCACTTCCCAACAGTAC 168  
QY 66 -----SerSerAsnSerProAlaIleValValLysProLysAlaLysGluSer 81  
DB 169 CAACCAACATCCATCCATCCACAGCCACCAACACCCCACTCAAACT 228  
QY 82 AsnThr----- 83  
DB 229 ACCACCATCACCACACACCGCCAGGAAACCAACCTCTCTCTGACACCAACCA 288  
QY 84 -----LysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAla 100  
DB 289 CCATTACCTCAAAAATGGAACTTCTCCAGAAACCGCTGCCACGGCTTGGACCTG 348  
QY 101 GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer 120  
DB 349 GAAACGGCGCTGCTCCGACAGCGCAACACCCCGCTCCCAACGGCGGACCCGAGG 408  
QY 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgAsnLeuPro 140  
DB 409 GTCCAAATCGCCGGAACTTCGCGCGGTGCCGAGCATGCGCGATCAAGACTCCCG 468  
QY 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160  
DB 469 GTGGTTCGGAATAATCCCAAAATGCAATTGACGGGTGTACGTGCGCAACGGTCCCAATCG 528  
QY 161 LeuHisGluProValThrGlyHisHisPheAspGlyAspGlyMetValHisAlaVal 180  
DB 529 CTCTAGACCTGTGGCCGGGACCACTTCTTCGACGGCGGCGCATGCTCCACCGCGTG 588  
QY 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200  
DB 589 AAGTTCACGAACGGCGCCAGCTAGCGCTCGCGCTTCACCGAGACGGAGCTCTCG 648  
QY 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyLysLeuHisGlyHis 220  
DB 649 CAGGAGAATCTCTAGCGCGCGCGGTGTTCCGAAAGCCCATCGGGAGCTCCACGGCCAC 708  
QY 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaGlyIleValAspPro 240  
DB 709 TCCGGCATCGCGCGCTCTCTCTACGGCGCGGTCTCTTCGGGCTCGTGTATGGG 768  
QY 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260  
DB 769 TCCACGGGATCGCGCGTGGCGAAGCGCGTCTCGTCTACTTCAACACCACTCTTGCC 828  
QY 261 MetSerGluAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280  
DB 829 ATGTCCGAAGACGATTTACCTTACCACGTGAGAATCACCCCTTAACGGGACCTTAACACC 888  
QY 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300  
DB 889 GTTGCCGCTTACGATTCACGGGCGAGCTCAACTCAACAAATGATCGCCACCCGAACTG 948  
QY 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeu 320  
DB 949 GACCCCGTCGACGGCGACCTCCACGGCTCAGCTACGAGCTCATTCAGNAGCTTACCTC 1008  
QY 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340  
DB 1009 AAGTACTTCCGTTCTCCCGCCGACGGCGTCAAGTCCCGCGAGCTGGAAATCCCGCTGAAG 1068  
QY 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValValProAspGln 360  
DB 1069 GAGCCACCATGATGACGATTCGCATTAACGGAGAAATTCGTGCTGCTCCCGACGAG 1128  
QY 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380  
DB 1129 CAGGTGGTCTCAACTACGAGAGATGATCACCGGGGGTCCCGGTGCTTACGACAAAG 1188  
QY 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400

DB 1189 AACAAAACTCAGGTTTGGGATTCTGCACAAAGATGCGAAGACCGGAATCGATGCGG 1248  
QY 401 TrpIleAspAlaProAspCysPheHisLeuTrpAsnAlaTrpGluGluProGlu 420  
DB 1249 TGGATCGACCGCCCGGATTTCTGCTTCCACTCTGGAACGCGTGGAGAGGCCGAA 1308  
QY 421 ThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440  
DB 1309 ACCGAGGAGGTGTGGTATGGTCTGTCATGACCCCTGCGGACTCCATTTTCAACGA 1368  
QY 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460  
DB 1369 TCGCAGGAGAGTTGAAGAGCGTGTGTGAGATGAAGCTGAACCTTGAGGACCGCAAG 1428  
QY 461 SerThrArgArgProIleLeuSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480  
DB 1429 TCCACTCGCGCGCCCATTTATCTCC---GACGCCGAACAAAGTGAACCTGGAAACCGCGCATG 1485  
QY 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500  
DB 1486 GTGAACAGAAACAAGCTCGGAAGGAAGACCCAGTTTCGGTATCTGGCTCTGGCGGAGCCC 1545  
QY 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis 520  
DB 1546 TGGCCCAAGTCTCGGGCTTTGCGAAAGTTGATTTGCTGAGTGGGAGAGTGAAGAAGTAC 1605  
QY 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGlyGlu 540  
DB 1606 ATGTATGAGAAAGAGAGTTGCGTGGGAGCGCTCTGTTTCTTCCC---AACGCCCAAA 1662  
QY 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560  
DB 1663 GAAGACGATGGTATATTTCTGCATTCTGCACGACGAGAAAGATGGAATCCGAGCTG 1722  
QY 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580  
DB 1723 CAGATTGTGAATGCCAAAATTTAAAGCTCGAAAGCTTCATCAAACTCCCTCTCGTGT 1782  
QY 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGln 597  
DB 1783 CCTACGGTTTTCATGGAACCTTTCATTCATCCAAAGATTTGAGGAACAA 1833  
RESULT 15  
BD017434  
LOCUS BD017434 1839 bp DNA linear PAT 27-AUG-2002  
DEFINITION Transgenic plant using neozanthine cleaving enzyme gene.  
ACCESSION BD017434  
VERSION BD017434.1 GI:22558610  
KEYWORDS JP 2001258579-A/6.  
SOURCE Vigna unguiculata (cowpea)  
ORGANISM Vigna unguiculata  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Vigna.  
REFERENCE 1 (bases 1 to 1839)  
AUTHORS Iuchi, K., Kobayashi, M. and Shinozaki, K.  
TITLE Transgenic plant using neozanthine cleaving enzyme gene  
JOURNAL Patent: JP 2001258579-A 6 25-SEP-2001;  
THE INSTITUTE OF PHYSICAL AND CHEMICAL RESEARCH  
COMMENT OS Vigna unguiculata  
PN JP 2001258579-A/6  
PD 25-SEP-2001  
PF 11-JAN-2001 JP 2001003476  
PI KIYOSHI IUCHI,MASATOMO KOBAYASHI,KAZUO SHINOZAKI PC  
C12N15/09,A01H5/00,C12N5/10,C12N5/02,C12N15/00,C12N5/00 CC  
Transgenic plant using neozanthine cleaving enzyme gene FH Key  
Location/Qualifiers (1) (1839).  
FT CDS Location/Qualifiers  
1..1839  
/organism="Vigna unguiculata"



/mol\_type="genomic DNA"  
/db\_xref="taxon:3917"  
BASE COUNT 447 a 595 c 442 g 355 t  
ORIGIN

Alignment Scores:  
Pred. No.: 178e-148 Length: 1839  
Score: 2168.50 Matches: 423  
Percent Similarity: 78.44% Conservative: 61  
Best Local Similarity: 68.56% Mismatches: 98  
Query Match: 68.84% Indels: 35  
DB: 6 Gaps: 8

US-09-758-269-6 (1-595) x BD017434 (1-1839)

Qy 8 AlaAlaValSerGlyArgTyrLeuGlyGlyAsnHisThrGlnProProLeuSerSerSer 27  
Db 7 TCATCAGCTTCAAAACACTTGTGTTAAACGCCACACTCCCATCTCCGCCCTTCAAAAGACCTA 66

Qy 28 GlnSerSerAspLeuSerTyrCysSerSerLeuProMetAlaSerArgValThrArgLys 47  
Db 67 CCTTCCACATCT--TCTCCCAAAACTTACTTCTTTA-----AGGAAACATCTCTCT 117

Qy 48 LeuAen---ValSerSerAlaLeuHisThrProProAlaLeuHisPheProLysGln--- 65  
Db 118 TCCACACCATCATCATGTTCCCTTCAACA-----CTCCACTTCCCAACACAGTAC 168

Qy 66 -----SerSerAsnSerProAlaLeuValValLysProLysAlaLysGluSer 81  
Db 169 CAACCAACATCCACATCCACATCCACAGCCACCACCAACACGCCCAATCAAAACT 228

Qy 82 AsnThr----- 83

Db 229 ACCACCATCACACACACCGCCCGAGGAACCAACCTCTCTCTGACCAACCA 288

Qy 84 -----LysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAla 100  
Db 289 CCATTACCTCAAAATGGAATCTTCCAGAAAGCGGTGCCAGCGCTTGACCTGGTC 348

Qy 101 GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer 120  
Db 349 GAAACGGCGCTCGTCTCGACGAGCGAAACACCGCTCCCAAAACGGCGGACCCGAGG 408

Qy 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuPro 140  
Db 409 GTCCAAATCGCGGGAATCTCGCGCGGTGCGGAGCATGCGCGGATCAAGGACTCCCG 468

Qy 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160  
Db 469 GTGGTCGGAATAATCCCAATGCATTGACGCGGTGTACGTGCGCAACGGTGCCTCCG 528

Qy 161 LeuHisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaVal 180  
Db 529 CTCTACGAGCTGTGCGCGCGGCACCATCTTTCGACGCGGCGCATGGTCCACGCGGTG 588

Qy 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200  
Db 589 AAGTTACGAACGGCGCGCGCATGCTACCGCTGCGCTTCCAGAGACGACGCTCTCTCG 648

Qy 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220  
Db 649 CAGGAGAAATCTTAGGCGCGCGGTGTTCCGAAGGCGCATCGGGGAGCTCCACGGCCAC 708

Qy 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaGlyIleValAspPro 240  
Db 709 TCCGCGATCGCGCGGTCTCTCTTCTACGCGCGGTCTCTTCCGCGCTCGTTGATGGG 768

Qy 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260  
Db 769 TCCAGGCGATGGGCGGTGGCAACCGCGTCTCGTCTACTTCAACCAACCACTCTTGCC 828

Qy 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280  
Db 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280

Search completed: November 15, 2003, 15:02:42

Job time : 6417 secs

Db 829 ATGTCGGAAGACGATTTTACCTACCACTGAGAAATACCCCTTAACGGCGACTTAACCAACC 888

Qy 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300  
Db 889 GTTGGCGGTTACGACTTCAACGGCGAGCTCAACCAATGATGCGGCCACCCGAAACTG 948

Qy 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValLysProTyrLeu 320  
Db 949 GACCCCGTCGACGGCGACTCCACGCGCTCAGCTACGACGCTCATTCAGAGCCTTACTC 1008

Qy 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340  
Db 1009 AAGTACTTCGTTTCTCCCGCAGCGGTCAAGTCCCGCGAGTGAATCCCTCTGAAG 1068

Qy 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGln 360  
Db 1069 GAGCCCAACCATGATGACAGATTTCCGCAATACGAGAAATTCGTCGTCGCCGACGAG 1128

Qy 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380  
Db 1129 CAGTGTGTTCTTCAAACTAACGAGATGATCACCGCGGTCCCCCGTGGTCTACGACAG 1188

Qy 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400  
Db 1189 AACAAAACCTCACGGTTTGGGATTTCTGCACAAGAAATCGAAGGACGCGATGCGGCG 1248

Qy 401 TrpIleAspAlaProAspCysPheHisLeuTrpAsnAlaTrpGluProGlu 420  
Db 1249 TGGATCAGCGCGCGGATGTTTCTGCTTCCACTCTGGAACGCGTGGGAGGACCCGAA 1308

Qy 421 ThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440  
Db 1309 ACCGAGAGGTTGGTGATGGTCTGATGATGACCTCGCGACTCCATTTTCAACGAA 1368

Qy 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460  
Db 1369 TGCAGGAGAGTTTGAAGAGCGTCTGTACAGATAAGGTGAACCTTGAGGACGGCAAG 1428

Qy 461 SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480  
Db 1429 TCCACTCGCGCGCCATTATCTCC--GACGCGCAACAAAGTGAACCTGGAGCCGGCATG 1485

Qy 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500  
Db 1486 GTGAAACAGAAACAAAGCTCGGAAGAACCCAGTTCGCGTATCTGGCTCTGGCGGAGCCC 1545

Qy 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis 520  
Db 1546 TGGCCCAAGTCTCGGCTTTGCGAAAGTTGATTTGCTGAGTGGGGAAGTGAAGAAGTAC 1605

Qy 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGlyGlyGlu 540  
Db 1606 ATGTATGAGAAAGAAAGTTCGGTGGGAGCCTCTGTTTCTTCC--AAACGGGCAAAA 1662

Qy 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560  
Db 1663 GAAGACATGGGTATATCTGGCATTCGTGCACACGAGAAAGAAATGGAATTCGAGCTG 1722

Qy 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580  
Db 1723 CAGATTGTGAATGCCCAAAATTTAAAGCTCGAAGCTTCCATCAAACTCCCTCTCTGTT 1782

Qy 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGln 597  
Db 1783 CCCTACGCTTTTCATGGAACCTTTCATTCATTCGAAGGATTTGAGGAACAA 1833



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 15, 2003, 09:19:43 ; Search time 447 Seconds  
(without alignments)  
3617.368 Million cell updates/sec

Title: US-09-758-269-6  
Perfect score: 3150  
Sequence: 1 MASFTATAAASGRWLGNGHT.....VPYGFHGTGIGADDLAKQVV 599

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cpn2.1/USPTO spool/US09758269/runat 14112003 135755 28735/app query.fasta\_1.775  
-DB=N Geneseq 15Jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORSE=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09758269 @CGN 1.1 312 @runat 14112003 135755 28735 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3150	100.0	1800	22	AAD09396	Arabidopsis thalia
2	2280.5	72.4	1818	22	AAD09401	Lycopersicon escul
3	2168.5	68.8	1839	22	AAD09399	Vigna unguiculata
4	1991	63.2	1752	22	AAD09394	Arabidopsis thalia
5	1930	61.3	1815	22	AAD09400	Zea mays neoxanthi
6	1663.5	52.8	1734	22	AAD09398	Arabidopsis thalia
7	969	30.8	1788	21	AAC42989	Arabidopsis thalia
8	969	30.8	1788	22	AAD09395	Arabidopsis thalia
9	969	30.8	1788	24	ABZ13639	Arabidopsis thalia
10	960.5	30.5	1950	22	AAF77206	cDNA encoding sunf
11	939	29.8	1777	21	AAC36083	Arabidopsis thalia
12	938	29.8	1617	22	AAD09397	Arabidopsis thalia
13	937	29.7	1617	22	AAD09402	Arabidopsis thalia
14	608	19.3	443	21	AAC56678	Eucalyptus grandis
15	592	18.8	492	22	AAC82706	Rice abscisic acid
16	556	17.7	491	21	AAC57157	Pinus radiata tran
17	522	16.6	393	25	ABX20484	Human GDP-mannose
18	511	16.2	372	21	AAC56695	Eucalyptus grandis
19	494.5	15.7	4403765	22	AAI99683	Mycobacterium tube
20	494.5	15.7	4411529	22	AAI99682	Mycobacterium tube
21	475.5	15.1	4403765	22	AAI99683	Mycobacterium tube
22	475.5	15.1	4411529	22	AAI99682	Mycobacterium tube
23	463	14.7	325	21	AAC56548	Eucalyptus grandis
24	415	13.2	386	21	AAC57162	Pinus radiata tran
25	409.5	13.0	340	21	AAC56520	Eucalyptus grandis
26	401	12.7	329	21	AAC57165	Pinus radiata tran
27	398	12.6	326	21	AAC57167	Pinus radiata tran
28	371.5	11.8	398	21	AAC57145	Pinus radiata tran
29	369	11.7	320	25	ABX20813	Human GDP-mannose
30	337.5	10.7	238	21	AAC57169	Pinus radiata tran
31	330.5	10.5	567	21	AAC56546	Eucalyptus grandis
32	322.5	10.2	1713	22	AAD19635	Arabidopsis thalia
33	322	10.2	470	24	ABL57537	Apoptosis inhibito
34	311	9.9	246	21	AAC57168	Pinus radiata tran
35	311	9.9	412	21	AAC57164	Pinus radiata tran
36	307.5	9.8	1791	22	AAD19634	Arabidopsis thalia
37	290	9.2	308	21	AAC57142	Pinus radiata tran
38	286.5	9.1	395	21	AAC57117	Pinus radiata tran
39	279	8.9	534	22	AAH44254	Physcomitrella pat
40	272.5	8.7	190	21	AAC57163	Pinus radiata tran
41	272.5	8.7	501	21	AAC43547	Zea mays DNA fragm
42	266.5	8.5	508	21	AAC56511	Eucalyptus grandis
43	260	8.3	3111	21	AAA53888	Beta, Beta-caroten
44	256	8.1	1773	22	AAH42170	Nucleotide sequenc
45	252.5	8.0	447	21	AAC56480	Eucalyptus grandis

ALIGNMENTS

RESULT 1  
AAD09396  
ID AAD09396 standard; cDNA; 1800 BP.  
AC AAD09396;  
XX  
XX  
XX 10-SEP-2001 (first entry)  
DE Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED3 cDNA.  
XX  
XX Neoxanthin cleavage enzyme; AtNCED3; abscisic acid; ABA; herbicide;  
KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
KW plant growth protectant; ss.  
XX  
XX Arabidopsis thaliana.  
XX  
XX  
XX Key Location/Qualifiers

```
FT CDS 1..1800
FT /*tag= a
FT /product= "Arabidopsis thaliana AtNCED3 protein"
XX
XX EP1116794-A2.
XX
XX 18-JUL-2001.
XX
XX 11-JAN-2001; 2001EP-0300218.
XX
XX 13-JAN-2000; 2000JP-0010056.
XX
XX 11-JAN-2001; 2001JP-0003476.
XX
XX (RIKE ) RIKEN KK.
XX
XX Iuchi S, Kobayashi M, Shinozaki K;
XX WPI; 2001-400081/43.
XX P-PSDB; RAE04784.
XX
XX A DNA encoding a protein with a neoxanthin cleavage activity for
XX producing transgenic plants with improved or decreased stress tolerance
XX
XX
XX Claim 3; Page 32-36; 101pp; English.
XX
XX The invention relates to neoxanthin cleavage enzymes and their
XX corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key
XX role in endogenous abscisic acid (ABA) biosynthesis under drought stress.
XX Neoxanthin cleavage enzyme is used for improving stress tolerance in a
XX plant when expressed in a plant cell. The invention also relates to
XX methods for increasing or decreasing stress tolerance in a plant by
XX introducing the DNA into the plant, and a transgenic plant into which a
XX neoxanthin cleavage enzyme is introduced. The improvement of stress
XX tolerance in plants is useful, for example in plant breeding. Neoxanthin
XX cleavage enzyme genes are useful for producing transgenic plants. An arid
XX land can be improved by growing transformat weed for several years and
XX then removing the weed by specifically lowering stress tolerance in the
XX weed by inducing an inducible promoter. The present cDNA sequence encodes
XX Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED3 protein.
XX The AtNCED3 cDNA is obtained from an Arabidopsis plant-derived cDNA
XX library using a cDNA of the CPRd65 (Cowpea Responsive to Dehydration)
XX gene isolated from cowpea plant as a probe.
XX
XX SQ Sequence 1800 BP; 458 A; 464 C; 439 G; 439 T; 0 other;

Alignment Scores:
Pred. No.: 3.92e-298 Length: 1800
Score: 3150.00 Matches: 599
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-758-269-6 (1-599) x AAD09396 (1-1800)

Qy 1 MetAlaSerPheThrAlaThAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThr 20
Db 1 ATGGCTCTTTTCACGGCAACGGCTCGCGTTTCTGGAGATGGCTTGGTGGCAATCATACT 60
Qy 21 GlnProLeuSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet 40
Db 61 CAGCGCCCATATGCTGCTCTCTCAAGCTCCGACTTGAGTTAGTCTCTTACCTATG 120
Qy 41 AlaSerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProAlaLeu 60
Db 121 GCCAGTCTGTGCACAGTAAGCTCAATGTTTCATCTGGCTTCACACTCTCCAGCTCT 180
Qy 61 HisPheProLysGlnSerSerAsnSerProAlaLeuValValLysProLysAlaLysGlu 80
Db 181 CATTTCCCTAAGCAATCATCAAACTCTCCCGCCATTGTTGTTAAGCCCAAGCCAAAGAA 240
Qy 81 SerAenThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaLeuAspAlaAla 100
```

```
Db 241 TCACAACACTAAACAGATGAATTTGTTCCAGAGAGCGCGCGGAGGTTGGAGCGCGCG 300
Qy 101 GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer 120
Db 301 GAGGGTTTCTTGTGACGACACGAGAGCTACACCGCTTCTTAAACACGGTGATCTCTAGT 360
Qy 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuPro 140
Db 361 GTTCAGATCCGCGAAATTTTGCTCCGGTGAATGAACAGACCCGTCGCGGCTAATCTTCCG 420
Qy 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160
Db 421 GTGGTCGAAAACCTTCCGATTCATCAAGAGAGTGATGTGCGCAACGAGCTAACCCA 480
Qy 161 LeuHisGluProValThrGlyHisHisPheAspGlyAspGlyMetValHisAlaVal 180
Db 481 CTTTCAGAGCGGTGACAGGTCACCACTTCTCGACGAGACGCTATGGTTTCACGCGGTC 540
Qy 181 LysPheGluHisGlySerHisSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200
Db 541 AAATTCCAAACACCGGTTCCAGTAGCTACGCTTGCCTGCGGTTTACTCAGACTAACCGGTTTGT 600
Qy 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220
Db 601 CAGGAAGCTCAATTGGGTCACCGGTTTCCCCAAGCCATCGTGAGCTTCACGGCCAC 660
Qy 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspPro 240
Db 661 ACCGTTATTCCCGACTCATGCTATTCTACGCCAGAGCTGCAGCGGTATAGTCGACCCG 720
Qy 241 AlaHisGlyThrGlyValAlaAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260
Db 721 GCACACGAAACCGGTAGCTAAGCGCGTTGGTCTATTTCATTAATGCGCGGTTATTGGCT 780
Qy 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280
Db 781 ATGTCGGAGAGATGATTACCTTACCAAGTTCCAGATCACTCCCAATGGAGATTAAAAACC 840
Qy 281 ValGlyArgPheAspPheAspGlyGlnLeuGlySerThrMetIleAlaHisProLysVal 300
Db 841 GTTGGTCGGTTTCGATTTTGTATGGACAAATTAGAAATCCAAATGATTGCCCCACCCGAAAGTC 900
Qy 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeu 320
Db 901 GACCCGGAATCCGGTGAACCTTTCGCTTAAAGCTTACACGCTCGTGTTCAAAGCCTTACCTA 960
Qy 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340
Db 961 AATATCTTCGATTTCTCACCAGCGGAACCTAATCACCAGGACGTCGAGATTACGTTGAT 1020
Qy 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValValProAspGln 360
Db 1021 CAGCCAAACGATGATGCACGATTTCCGATTTACAGAGAACTTCGTCGTCTGACCTGACCCAG 1080
Qy 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380
Db 1081 CAAGTCGTTTTCAGCTGCCGAGATGATCCGCGGTGGGTCTCCGGTGGTTTACGACAAG 1140
Qy 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400
Db 1141 AACAAAGTCGCAAGATTCGGGATTTTAGACAAATACCGCGAGATTTCATCGAACATTAAAG 1200
Qy 401 TrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluProGlu 420
Db 1201 TGGATTGATGTCCTCCAGATTGCTTCTGCTTCCATCTCTGGAACGCTTGGGAAGAGCCAGAA 1260
Qy 421 ThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440
Db 1261 ACAGATGAAGTCGCTGATAGGTCCTGTATGACTCCACAGACTCAATTTTCAACGAG 1320
Qy 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460
```

```
Db 1321 TCTGACGAGAAATCTCAAGAGTGTCTGTCTGAAATCCGCTGAATCTCAAAACCGGTGAA 1380
Qy 461 SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480
Db 1381 TCAACTCGCGCTCCGATCATCTCCAAACGAAGATCAACAAGTCAACCTCGAAGCAGGATG 1440
Qy 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500
Db 1441 GTCAACAGAAACATGCTCGCGCGTAAACCAAAATTCGCTTACTTGCGCTTTAGCCGAGCG 1500
Qy 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis 520
Db 1501 TGGCCTAAAGTCTCAGGATTCGTAAGTTGATCTCAGTCTACGAGAGTGAAGTGAAGAAAT 1560
Qy 521 LeuTyrGlyAspAsnArgTyrGlyGluGluProLeuPheLeuProGlyGlyGlyGlu 540
Db 1561 CTTTACGGCGATAACCGTTACGAGGAGGAGCGCTCTGTTCTCCCGAGAGGAGAG 1620
Qy 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560
Db 1621 GAAGACGAAGAGATACATCTCTGTTCTCGTTCACGACGAGAGACATCGGAATTCGGAGTTA 1680
Qy 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580
Db 1681 CAGATAGTTAACCGCTTAGCTTAGAGTTGAAGCAACGGTTAAACTTCCGTCAGGGTT 1740
Qy 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGlnValVal 599
Db 1741 CCGTACGGATTTTCACGGTACATTCATCGGAGCGGATGATTTGCGAAGCAGGTCGTG 1797

RESULT 2
AAD09401
ID AAD09401 standard; cDNA; 1818 BP.
XX
AC AAD09401;
XX
DT 10-SEP-2001 (first entry)
XX
DE Lycopersicon esculentum neoxanthin cleavage enzyme, LeNCE1 cDNA.
XX
KW Tomato; neoxanthin cleavage enzyme; LeNCE1; abecisic acid; ABA;
KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;
KW plant growth protectant; herbicide; ss.
XX
OS Lycopersicon esculentum.
XX
FH Key Location/Qualifiers
FT CDS 1..1818
FT /tag= a
FT /product= "Lycopersicon esculentum LeNCE1 protein"
XX
XX EP1116794-A2.
XX
XX 18-JUL-2001.
XX
XX 11-JAN-2001; 2001EP-0300218.
XX
XX 13-JAN-2000; 2000JP-0010056.
XX
XX 11-JAN-2001; 2001JP-0003476.
XX
XX (RIKE ) RIKEN KK.
XX
XX Iuchi S, Kobayashi M, Shinozaki K;
XX
XX WPI: 2001-400081/43.
XX
XX P-PSDB; AAE04789.
XX
XX A DNA encoding a protein with a neoxanthin cleavage activity for
XX producing transgenic plants with improved or decreased stress tolerance
XX
XX Claim 3; Page 67-71; 101pp; English.
XX
XX
```

```
CC The invention relates to neoxanthin cleavage enzymes and their
CC corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key
CC role in endogenous abscisic acid (ABA) biosynthesis under drought stress.
CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a
CC plant when expressed in a plant cell. The invention also relates to
CC methods for increasing or decreasing stress tolerance in a plant by
CC introducing the DNA into the plant, and a transgenic plant into which a
CC neoxanthin cleavage enzyme is introduced. The improvement of stress
CC tolerance in plants is useful, for example in plant breeding. Neoxanthin
CC cleavage enzyme genes are useful for producing transgenic plants. An arid
CC land can be improved by growing transformant weed for several years and
CC then removing the weed by specifically lowering stress tolerance in the
CC weed by inducing an inducible promoter. The present cDNA sequence encodes
CC Lycopersicon esculentum neoxanthin cleavage enzyme, LeNCE1 protein
CC related to the invention.
XX
SQ Sequence 1818 BP; 569 A; 383 C; 369 G; 497 T; 0 other;

Alignment Scores:
Pred. No.: 5,17e-213 Length: 1818
Score: 2280.50 Matches: 435
Percent Similarity: 82.04% Conservative: 63
Best Local Similarity: 71.66% Mismatches: 90
Query Match: 72.40% Indels: 19
DB: 22 Gaps: 6

US-09-758-269-6 (1-599) x AAD09401 (1-1818)
Qy 5 ThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThrGlnProLeu 24
Db 7 ACTACTACTTCCATGCGCCAAATACATACATGATTTTTCGATCAAACTCTATTTCTCTACTCAAAAT 51
Qy 25 SerSerSerGlnSerSerAspLeuSerTyrCysSer-----SerLeuProMetAla 41
Db 52 TCATATGCCATCATCAAGGAGTGTGGTTTTCGATCAAACTCTATTTCTCTACTCAAAAT 111
Qy 42 SerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProProAlaLeuHis 61
Db 112 CAACATAATAGGCAAAAGTCTCAACATTAATTCCTCTCTTCAAGTCTCCACCTATATCTCAT 171
Qy 62 PheProLysGlnSerSerAsn-----SerProAla-----IleValValLysProLys 77
Db 172 TTTCTTAACCAATCTTCAAAATTTATCAACACCAAGAAATATACAAATTTTCACACCCAAA 231
Qy 78 AlaLysGluSerAsn-----ThrLysGlnMetAsnLeuPheGlnArgAla 92
Db 232 CAAGAAACAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 291
Qy 93 AlaAlaAlaLeuAspAlaAlaGluGlyPheLeuValSerHisGluLysLeuHisPro 112
Db 292 GCAGCAATGGCTTTAGATGCTGTAGAAAGTCTTTAACTAAACATGAACCTTGACACCCCT 351
Qy 113 LeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValAsnGlu 132
Db 352 TTGCGGAAACAGCGCCACCCACGAGTCCAGATTTCTGGGAAATTTTGCTCCGCTACCGGAA 411
Qy 133 GlnProValArgAsnLeuProValValGlyLysLeuProAspSerIleLysGlyVal 152
Db 412 AATCAGTCTGTCATCTCTCCGTCACCGGAAATATACCAATGTGTTCAGGCGTT 471
Qy 153 TyrValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHisHisPhePheAsp 172
Db 472 TAGCTTCGAAACGGAGCTAACCTCTTTTGAACCAACCGCGGACACCATTTCTTCGAC 531
Qy 173 GlyAspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyrAlaCysArg 192
Db 532 GGCAGCGTATGTTACGCGCTTCAATTTCAAAATGGGTGCGCTAGTTACGCTTCCGCT 591
Qy 193 PheThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPheProLys 212
Db 592 TTCACCTGAACAGAGAGGCTTTGTTCAAGAAAGAAAGCTTTGGGTGCGCCCTGTTTCCCTAAA 651
Qy 213 AlaIleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArg 232
```

Db 652 GCCATTGGTGAATACATGCTCTGGAATTCGAAGCTTATGCTGTTTTCACGCTCGT 711  
Qy 233 AlaAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuVal 252  
Db 712 GGGCTCTTCGGACTTGTGATCACAGTAAGAACTGCTGTGTGCAAAACGCCGGTTAGTC 771  
Qy 253 TyrPheAsnGlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIle 272  
Db 772 TATTTCAATAACCGGATACCTGCTATGCTGAAGATGATTTGCCTTACCATGTAAAGGTA 831  
Qy 273 ThrProAsnGlyAspLeuLeuThrValGlyArgPheAspPheAspGlyGlnLeuGluSer 292  
Db 832 ACACCCACCGCGGATCTTAAACACAGAGGTTCGATTCGATTCGACGCCGCTAAATCC 891  
Qy 293 ThrMetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeuSerTyr 312  
Db 892 ACCATGATAGCTACCCCAAGCTCGACCCAGTTTCCGGTGAGCTATTGCTCTTAGCTAC 951  
Qy 313 AspValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSer 332  
Db 952 GATGTGATTCAGAAGCCATACCTCAAGTACTTCAGATTTTCAAAAAATGGGAAAAATCA 1011  
Qy 333 ProAspValGluIleGlnLeuAspGlnProThrMetHisAspPheAlaIleThrGlu 352  
Db 1012 AATGATGTTGAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 1071  
Qy 353 AsnPheValValProAspGlnGlnValValPheLysLeuProGluMetIleArgGly 372  
Db 1072 AACTTCGTCGTCATTCCTGATCAACAAGTCGTTTCAAGATGTCGAAATGATCCGTGGA 1131  
Qy 373 GlySerProValValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyr 392  
Db 1132 GGTTACCCGGTGGTTTACGACAAAGAACTTCCCGATTTGGTATTTCTGGATTAAGTAC 1191  
Qy 393 AlaGluAspSerAsnIleLysTyrIleAspAlaProAspCysPheCysPheHisLeu 412  
Db 1192 CGGAAGATGGTCTGATTTGAATGGTGGTGAAGTACCTGATTTGTTCTGTTCCACTC 1251  
Qy 413 TrpAsnAlaTrpGluGluProGluThrAspGluValValIleGlySerCysMetThr 432  
Db 1252 TGGATGCTTGGGAAGACGAGAACACATGAAATCGTTGTAATTTGGTTTCATGTATGACA 1311  
Qy 433 ProProAspSerIlePheAsnGluSerAspGluAsnLysSerValLeuSerGluIle 452  
Db 1312 CCACCAGACTCCATTTCAATGAATGTGATGAAGGGCTAAAGAGTGTGTTATCCGAATC 1371  
Qy 453 ArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGln 472  
Db 1372 CGTCTCAATTTGAAACAGGGAATCAACAGAAATCCATATCGAAACCCCGATGAA 1431  
Qy 473 GlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPhe 492  
Db 1432 CAAGTGAATTTAGAGCTGGAATGGTGAACCGAAACAACTCGGAAGGAAACAGAGTAT 1491  
Qy 493 AlaTyrLeuAlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAspLeu 512  
Db 1492 GCTTATTTGGCTATCGCTGAACCATCGGCAAAAGTTCTGTTTTGCAAAAGTAAACCTG 1551  
Qy 513 ThrThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGluProLeu 532  
Db 1552 TTCACCGGTGAAGTTGAGAAATTCATTTATGTTGACAAACAAATATGTTGGGGAACCTCTT 1611  
Qy 533 PheLeuProGlyGlu-----GlyGlyGluGluAspGluGlyTyrIleLeuCysPheVal 550  
Db 1612 TTTTACCAGAGACCCCAACAGCAAGGAAGACGATGTTATATTTAGCTTCGTT 1671  
Qy 551 HisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaValSerLeuGluVal 570  
Db 1672 CACGATGAGAAGAAATGGAATTCAGAACTGCAAAATGTTTAAACGCAATGAGTTTGAAGTTG 1731  
Qy 571 GluAlaThrValLysLeuProSerArgValProTyrGlyPheHisGlyThrPheIleGly 590

Db 1732 GAGCAACTGTGAAGCTTCATCAAGAGTTCTTATGGAATTCATGGAACATCATAAAC 1791  
Qy 591 AlaAspAspLeuAlaLysGln 597  
Db 1792 GCCAATGATTGGCAAAATCAG 1812  
RESULT 3  
ID AAD09399 standard; cDNA; 1839 BP.  
XX AAD09399;  
AC AAD09399;  
XX 10-SEP-2001 (first entry)  
XX Vigna unguiculata neoxanthin cleavage enzyme, CPRD65 cDNA.  
DE Cowpea; neoxanthin cleavage enzyme; abscisic acid; ABA; herbicide;  
XX stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
KW plant growth protectant; CowPea Responsive to Dehydration; CPRD65; ss.  
XX Vigna unguiculata.  
OS  
FH Key Location/Qualifiers  
FT CDS 1..1839  
FT /+tag= a  
FT /product= "Vigna unguiculata CPRD65 protein"  
XX EPI116794-A2.  
XX 18-JUL-2001.  
XX 11-JAN-2001; 2001EP-0300218.  
XX 13-JAN-2000; 2000JP-0010056.  
PR 11-JAN-2001; 2001JP-0003476.  
XX (RIKE ) RIKEN KK.  
XX Iuchi S, Kobayashi M, Shinozaki K;  
XX WPI; 2001-400081/43.  
DR P-PSDB; AAE04787.  
XX  
PT A DNA encoding a protein with a neoxanthin cleavage activity for  
FT producing transgenic plants with improved or decreased stress tolerance  
PT  
XX  
PS Claim 3; Page 53-56; 101pp; English.  
XX  
CC The invention relates to neoxanthin cleavage enzymes and their  
CC corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key  
CC role in endogenous abscisic acid (ABA) biosynthesis under drought stress.  
CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a  
CC plant when expressed in a plant cell. The invention also relates to  
CC methods for increasing or decreasing stress tolerance in a plant by  
CC introducing the DNA into the plant, and a transgenic plant into which a  
CC neoxanthin cleavage enzyme is introduced. The improvement of stress  
CC tolerance in plants is useful, for example in plant breeding. Neoxanthin  
CC cleavage enzyme genes are useful for producing transgenic plants. An arid  
CC land can be improved by growing transformant weed for several years and  
CC then removing the weed by specifically lowering stress tolerance in the  
CC weed by inducing an inducible promoter. The present cDNA sequence encodes  
CC Vigna unguiculata neoxanthin cleavage enzyme, CPRD65 (CowPea Responsive  
CC to Dehydration) protein. CPRD65 gene is isolated from cowpea plant.  
XX  
SQ Sequence 1839 BP; 447 A; 595 C; 442 G; 355 T; 0 other;  
Alignment Scores:  
Pred. No.: 4.84e-202 Length: 1839  
Score: 2168.50 Matches: 423  
Percent Similarity: 78.44% Conservative: 61  
Best Local Similarity: 68.56% Mismatches: 98  
Query Match: 68.84% Indels: 35

DB: 22 Gaps: 8

US-09-758-269-6 (1-599) x AAD09399 (1-1839)

QY 8 AlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThrGlnProProLeuSerSer 27  
 DB : : : : :  
 DB 7 TCATCAGCTTCAAAACACTTGGTTTAACGCCACACTCCCATCTCCCCCTTCAAAACCTA 66  
 QY 28 GlnSerSerAspLeuSerTyrCysSerSerLeuProMetAlaSerArgValThrArgLys 47  
 DB : : : : :  
 DB 67 CCTTCACATCT--TCTCCCAAACTTACTTCTCTTA-----AGGAAACATCCTCT 117  
 QY 48 LeuAsn---ValSerSerAlaLeuHisThrProAlaLeuHisPheProLysGln--- 65  
 DB : : : : :  
 DB 118 TCCAACACCATCACAATGTTCCCTTCAACA-----CTCCACTTCCCAACAGTAC 168  
 QY 66 -----SerSerAsnSerProAlaIleValValLysProLysAlaLysGluSer 81  
 DB : : : : :  
 DB 169 CAACCAACATCCACATCCACATCCACAGCCACCAACACACCCCAATCAAACT 228  
 QY 82 AsnThr----- 83  
 DB : : : : :  
 DB 229 ACCACCATCACCACCAACACCGCCAGGAAACCAACCTCTCTCTGACACCAACCA 288  
 QY 84 -----LysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAla 100  
 DB : : : : :  
 DB 289 CCAATACCTCAAAATGGAACTTTCTCCAGAAAGCGCTGCCACGGCTTGGACCTGTC 348  
 QY 101 GluGlyPheLeuValSerHisGlyLysLeuHisProLeuProLysThrAlaAspProSer 120  
 DB : : : : :  
 DB 349 GAAACGGCGCTCTCTCGCACGAGCGAACAACCCGCTCCCAACAGCGGACCGGAGG 408  
 QY 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgAsnLeuPro 140  
 DB : : : : :  
 DB 409 GTCCAAATCGCCGGAACTTCGCGCGGTCCGCGAGCATGCCGCGATCAAGACTCCCG 468  
 QY 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160  
 DB : : : : :  
 DB 469 GTGGTCGGAATAATCCCAATATGCAATGACGCGGTGACGCGCAACGGTGCCAAATCG 528  
 QY 161 LeuHisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaVal 180  
 DB : : : : :  
 DB 529 CTCTACGAGCTTGCGCGGCGCACCTTCTTCGACGCGCGGATGTCACCGCGGTG 588  
 QY 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200  
 DB : : : : :  
 DB 589 AAGTTCAGAACGCGCGCCAGCTACGCTGCCGCTTCACGACGACGAGCTCTCTCG 648  
 QY 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220  
 DB : : : : :  
 DB 649 CAGGAGAAATCTCTAGCGCGCGCGGTGTTCCGAAAGGCGCATCGGGGAGCTCCACGCGCAC 708  
 QY 221 ThrGlyAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspPro 240  
 DB : : : : :  
 DB 709 TCCGCGATCGCGCGCTCTCTCTTCACGCGCGCGGTCTCTTCGCGGTCTGATGGG 768  
 QY 241 AlaHisGlyThrGlyValAlaAlaAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260  
 DB : : : : :  
 DB 769 TCCAGGCGATCGGGGTGGGAAACCGCGGTCTCTCTCTACTTCAACCAACCACTCTTGCC 828  
 QY 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280  
 DB : : : : :  
 DB 829 ATGTCGGAAGACGATTTACCTTACACGCGGAGCTCAACTCAACAATGATCGCCACCGCAACTG 888  
 QY 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300  
 DB : : : : :  
 DB 889 GTTGCGCGTACGATTCACGCGGAGCTCAACTCAACAATGATCGCCACCGCAACTG 948  
 QY 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeu 320  
 DB : : : : :  
 DB 949 GACCCCGTCGAGCGGACCTCCACGCGCTCAGCTACGAGCTCATTCAGAGCGCTTACCTC 1008  
 QY 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340

DB : : : : :  
 1009 AAGTACTTCCGTTTCTCCCGACGCGCTCAAGTCCCGCGAGCTGGAATCCCCCTGAAG 1068  
 QY 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGln 360  
 DB : : : : :  
 1069 GAGCCACCATGATGACAGATTTCCCATACGGAATTTCTGCTGCTCCCGACCAAG 1128  
 QY 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380  
 DB : : : : :  
 1129 CAGGTGGTCTTCAAACTAACCGAGATGATACCGCGCGGTCCCGCTGCTTACGACAAG 1188  
 QY 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400  
 DB : : : : :  
 1189 AACAAACCTCACGGTTTGGGATTTCTGCACAGAATGCGAAGGACCGAAATCGATGCGG 1248  
 QY 401 TrpIleAspAlaProAspCysPheHisLeuTrpAsnAlaTrpGluLysProGlu 420  
 DB : : : : :  
 1249 TGGATCGACGCGCGGATTTCTGCTTCCACCTCTGGAACGCTGGAGAGGCCGAA 1308  
 QY 421 ThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440  
 DB : : : : :  
 1309 ACCGAGGAGGTGGTGTGATTTGGTCTCTGATGACCCCTGCGGACTCCATTTTCAACGAA 1368  
 QY 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460  
 DB : : : : :  
 1369 TCGGAGGAGAGTTGAAGAGCGTGTGTGACAGATTAAGCTCAACTTGAGGACCGGCAAG 1428  
 QY 461 SerThrArgArgProIleIleSerAsnGluAspGlnValAsnLeuGluAlaGlyMet 480  
 DB : : : : :  
 1429 TCCACTCGCGCGCCCATTTATCTCC---GACCGCCCAACGAAGTGAACCTGGAAGCGCGCATG 1485  
 QY 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500  
 DB : : : : :  
 1486 GTGAACAAACAAGCTCGGAAGGAGAACCCAGTTCCGCTATCTGGCTCTGCGGAGCCC 1545  
 QY 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis 520  
 DB : : : : :  
 1546 TGGCCCAAGTCTCGGGCTTTGCGAAAGTTGATTTGCTAGTGGGGAAGTGAAGAAGTAC 1605  
 QY 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGlyGlyGlu 540  
 DB : : : : :  
 1606 ATGTATGAGAAGAGAGTTCGGTGGGAGCTCTGTTTCTTCCC---AACGCGCAAAAC 1662  
 QY 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560  
 DB : : : : :  
 1663 GAAGCATGAGTATATTTCTGCATTCGTGCACGACGAGAAGAATGGAATCCGAGCTG 1722  
 QY 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580  
 DB : : : : :  
 1723 CAGATTGTGAATGCCAAAATTTAAAGCTCGAAGCTTCCATCAAACTCCCTCTCGTGT 1782  
 QY 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspLeuAlaLysGln 597  
 DB : : : : :  
 1783 CCTACGCTTTTTCATGGAACCTTTCATTCATTCCAAGGATTTGAGAAACNA 1833

RESULT 4  
 AAD09394  
 ID AAD09394 standard; cDNA; 1752 BP.  
 XX  
 AC AAD09394;  
 XX  
 DT 10-SEP-2001 (first entry)  
 XX  
 DE Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED1 cDNA.  
 XX  
 KW Neoxanthin cleavage enzyme; AtNCED1; abscisic acid; ABA; herbicide;  
 KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
 KW plant growth protectant; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1752

```
/*tag= a
/product= "Arabidopsis thaliana AtNCED1 protein"
```

```

FT  EP1116794-A2.
XX  18-JUL-2001.
XX  11-JAN-2001; 2001EP-0300218.
XX  13-JAN-2000; 2000JP-0010056.
PR  11-JAN-2001; 2001JP-0003476.
XX  (RIKE ) RIKEN KK.
XX
XX  Iuchi S, Kobayashi M, Shinozaki K;
XX  WPI; 2001-400081/43.
DR  P-PSDB; AAE04782.
XX
XX  A DNA encoding a protein with a neoxanthin cleavage activity for
XX  producing transgenic plants with improved or decreased stress tolerance
XX
XX  Claim 3; Page 18-22; 101pp; English.
XX
XX  The invention relates to neoxanthin cleavage enzymes and their
XX  corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key
XX  role in endogenous abscisic acid (ABA) biosynthesis under drought stress.
XX  Neoxanthin cleavage enzyme is used for improving stress tolerance in a
XX  plant when expressed in a plant cell. The invention also relates to
XX  methods for increasing or decreasing stress tolerance in a plant by
XX  introducing the DNA into the plant, and a transgenic plant into which a
XX  neoxanthin cleavage enzyme is introduced. The improvement of stress
XX  tolerance in plants is useful, for example in plant breeding. Neoxanthin
XX  cleavage enzyme genes are useful for producing transgenic plants. An arid
XX  land can be improved by growing transformant weed for several years and
XX  then removing the weed by specifically lowering stress tolerance in the
XX  weed by inducing an inducible promoter. The present cDNA sequence encodes
XX  Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED1 protein.
XX  The AtNCED1 cDNA is obtained from an Arabidopsis plant-derived cDNA
XX  library using a cDNA of the CPRD5 (CowPea Responsive to Dehydration)
XX  gene isolated from cowpea plant as a probe.
XX
XX  Sequence 1752 BP; 484 A; 396 C; 428 G; 444 T; 0 other;
XX
XX  Alignment Scores:
XX  Pred. No.: 1.07e-194 Length: 1752
XX  Score: 1991.00 Matches: 380
XX  Percent Similarity: 78.44% Conservative: 82
XX  Best Local Similarity: 64.52% Mismatches: 105
XX  Query Match: 63.21% Indels: 22
XX  DB: 22 Gaps: 6
XX
XX  US-09-758-269-6 (1-599) x AAD09394 (1-1752)
XX
XX  16 GlyGlyAenHisThrGlnProLeuSerSerSerGlnSerSerAspLeuSerTyr--- 34
XX  31 GGTGGTATTAAACATGCGCCT-----CAAGCCCAAAATTGATTGGGTATTAGG 78
XX
XX  35 ---CysSerSerLeuProMetAlaSerArgValThrArgLysLeuAsnValSerSerAla 53
XX  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX  79 CCCATTAAAGACCAACCGAAGTTATTAAATGCACGGTGCAGATCGAGTAAACGAA--- 135
XX
XX  54 LeuHisThrProProAlaLeuHisPheProLysGlnSerSerAsnSerProAlaIleVal 73
XX  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX  136 TTAACCAAAAAACGCAATTTATTACACCCAGAACCCAGCTACTCCCGCG- 186
XX
XX  74 ValLysProLysAlaLysGluSerAsnThrLysGlnMetAsnLeuPheGlnArgAla 93
XX  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX  187 -----CAGCATTAATCCTCTCCGGCTAAACATCTCCAGAAACGCGCG 228
XX
XX  94 AlaAlaLeuAspAlaAlaGluGlyPheLeuValSerHisGluLysLeuHisProLeu 113
XX  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db  229 GCGATTCCGATCGACGCGCTGAGCGTGCATTAATCTCACACGAGCAAGATTCTCCACTT 288
Qy  114 ProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValAsnGluGln 133
Db  289 CCCAAAACCGCTGATCCACGTTTCAGATTCCCGGGAATATTATCCCGGATCCGGAATCT 348
Qy  134 ProValArgArgAsnLeuProValValGlyLysLeuProAspSerIleLysGlyValTyr 153
Db  349 TCCGTCGCGGAAACCTCACCGTCGAGGAACAATCCCTGACTGCTGCTGCGGTGTTAT 408
Qy  154 ValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHisHisPhePheAspGly 173
Db  409 ATCCGTAACGCGCGAATCCGATGTTTGAGCAACAGCTGGGCACCATTTATTTCGACGA 468
Qy  174 AspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyrAlaCysArgPhe 193
Db  469 GACGGAATGGTTTCACGCAGTTAAATAACCAACGCTTACGCTAGCTACGATGCGGTTT 528
Qy  194 ThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPheProLysAla 213
Db  529 ACAAAAACCGAGAGATTGGTTTCAGAAAACAGATTGGTCGACCAAGTTTCCCGAAGCA 588
Qy  214 IleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAla 233
Db  589 ATCGCGAGCTTCACGCTCACTCGGAATCGCAGTTTGATGCTGTTTACGCACTGGG 648
Qy  234 AlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyr 253
Db  649 CTTTGTGCTGTGATCAACAAACCGGCTCGGAGTAGCAACACGCGGTTTGGTTTAC 708
Qy  254 PheAsnGlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIleThr 273
Db  709 TTTAATAACCGGCTTTTAGTATGTTCAGAGACAGATTACCGTACCAATTAATAATTA 768
Qy  274 ProAsnGlyAspLeuLysThrValGlyArgPheAspPheAspGlyGlnLeuGluSerThr 293
Db  769 CAACACCGCGATCTCCAAACCGTTGGACGTTACGATTTCCAGCGTCAGTTAAAAATCC 828
Qy  294 MetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeuSerTyrAsp 313
Db  829 ATGATAGCTCACCCGAAACCTGGACCGGTTACGAAGAGAGCTTCACGCGTTAAGTAC 888
Qy  314 ValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSerPro 333
Db  889 GTCGTTAAGAAACCTTACCTGAAATACTTCAGATTCTCGCCAGACGCGGTTAATCC 948
Qy  334 AspValGluIleGlnLeuAspGlnProThrMetHisAspPheAlaIleThrGluAsn 353
Db  949 GAATTGGAGATCCCGCTCGAAACCTCCGACGATGATTCCAGATTTCGCTATATAACGG 1008
Qy  354 PheValValValProAspGlnValValPheLysLeuProGluMetIleArgGlyGly 373
Db  1009 TTTGTGTGATTCTCTGATCAACAAGTCGTGTTCAAGCTCGCGGAGAGATTTCCTGGT 1068
Qy  374 SerProValValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAla 393
Db  1069 TCTCCGGTTGTTTCGACGCGGAAAGGTTTCCGATGGGATAATGCCCAAGACGCG 1128
Qy  394 GluAspSerSerAsnIleLysTrpIleAspAlaProAspCysPheCysPheHisLeuTrp 413
Db  1129 ACAGAAGCTTCTCAGATAATCTGGGTGAACCTCTCCGAGACGCTTCTGTTTTCATCT 1188
Qy  414 AsnAlaTrpGluLupProGluThrAspGluValValIleGlySerCysMetThrPro 433
Db  1189 AATCATGTTGGAATCCCGGAGACGAGAGAGATTGTTGTTGATCGGATCGTGTATCTG 1248
Qy  434 ProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArg 453
Db  1249 GCGGATTCAATCTTCAACGAGAGACGAGAGCTTGAGACGCTTTGTGCGAGATCAGG 1308
Qy  454 LeuAsnLeuLysThrGlyGluSerThrArgArgProIleLeuSerAsnGluAspGlnGln 473
Db  1309 ATAAACCTCAGAACACGTAACCAACCGCTCGTTTCGTTTGGTTTAAACGAGGAT----- 1362

```

QY 474 ValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgIysThrIysPheAla 493  
 DB 1363 GTAAATTTAGATTGGTATGTTAAACCGAATCGGTTAGGAAGAAACCCGGTTCGGC 1422  
 QY 494 TyrLeuAlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAspLeuThr 513  
 DB 1423 TTTTGGCTATTGCTTATCTTGGCCAAAGTTTCGGTTTCGTAAGTTCGATCTTGC 1482  
 QY 514 ThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGluProLeuPhe 533  
 DB 1483 ACCGGTGCAGATGAATAATATTTACGGCGGTGAGAAATATGCGCGCAACCGTTTTC 1542  
 QY 534 LeuProGlyGluGlyGly-----GluGluAspGluGlyTyrIleLeuCysPhe 549  
 DB 1543 TTCCCGCGCAACTCCCGTAACGCGGAAGAAATGAAGATGACGGTTATATATTGTTCAC 1602  
 QY 550 ValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaValSerLeuGlu 569  
 DB 1603 GTTCATGACGAAGAAACAAAGACATCAGAGCTTCAGATTATTAAACGCTGTTAATTAAAG 1662  
 QY 570 ValGluAlaThrValLysLeuProSerArgValProTyrGlyPheHisGlyThrPheIle 589  
 DB 1663 CTTGAAGCTACGATTAAACTACCGTCTAGATACCGTATGGGTTTCATGGCAATTTG 1722  
 QY 590 GlyAlaAspAspLeuAlaLysGlnVal 598  
 DB 1723 GATTCCGAATGAATCGTTGATCAATTA 1749

## RESULT 5

AAD09400

ID AAD09400 standard; cDNA; 1815 BP.

XX

AC AAD09400;

DT 10-SEP-2001 (first entry)

XX

DE Zea mays neoxanthin cleavage enzyme, VP14 cDNA.

XX

KW Maize; neoxanthin cleavage enzyme; VP14; abscisic acid; ABA; herbicide;  
 KW stress tolerance; transgenic plant; plant breeding; antisense therapy;  
 KW plant growth protectant; ss.

XX

OS Zea mays.

XX

FH Key Location/Qualifiers

FT CDS 1..1815

FT /tag= a

FT /product= "Zea mays VP14 protein"

XX

PN EP1116794-A2.

XX

PD 18-JUL-2001.

XX

PF 11-JAN-2001; 2001EP-0300218.

XX

PR 13-JAN-2000; 2000JP-0010056.

XX

PR 11-JAN-2001; 2001JP-0003476.

XX

PA (RIKE ) RIKEN KK.

XX

PI Iuchi S, Kobayaashi M, Shinozaki K;

XX

DR WPI; 2001-400081/43.

XX

DR P-PSDB; AAE04788.

XX

PT A DNA encoding a protein with a neoxanthin cleavage activity for  
 PT producing transgenic plants with improved or decreased stress tolerance

XX

PS Claim 3; Page 60-64; 101pp; English.

XX

CC The invention relates to neoxanthin cleavage enzymes and their

CC corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key  
 CC role in endogenous abscisic acid (ABA) biosynthesis under drought stress.  
 CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a  
 CC plant when expressed in a plant cell. The invention also relates to  
 CC methods for increasing or decreasing stress tolerance in a plant by  
 CC introducing the DNA into the plant, and a transgenic plant into which a  
 CC neoxanthin cleavage enzyme is introduced. The improvement of stress  
 CC tolerance in plants is useful, for example in plant breeding. Neoxanthin  
 CC cleavage enzyme genes are useful for producing transgenic plants. An arid  
 CC land can be improved by growing transformant weed for several years and  
 CC then removing the weed by specifically lowering stress tolerance in the  
 CC weed by inducing an inducible promoter. The present cDNA sequence encodes  
 CC Zea mays neoxanthin cleavage enzyme, VP14 protein related to the  
 CC invention.

XX

SQ Sequence 1815 BP; 266 A; 689 C; 604 G; 256 T; 0 other;

## Alignment Scores:

Pred. No.: 1.05e-178 Length: 1815  
 Score: 1930.00 Matches: 375  
 Percent Similarity: 75.68% Conservative: 73  
 Best Local Similarity: 63.34% Mismatches: 122  
 Query Match: 61.27% Indels: 22  
 DB: 22 Gaps: 8

US-09-758-269-6 (1-599) x AAD09400 (1-1815)

QY 22 ProProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMetAla 41  
 DB 49 CCGGCCCGGTCAGGGCCCGGGCTCC-----AATTCGTAGGTTCTCG 93  
 QY 42 SerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrPro-----Pro 58  
 DB 94 CCGCGCGCGTCAGCTCGCGCCCGCGAGTCCTCCAGGCGCGTTCACAAGCCC 153  
 QY 59 AlalaLeuHisPheProLysGlnSerSerAsnSerProAlaIleValVal-----74  
 DB 154 GTCCGCGACCTGCCTCGCGCGCTCCAGGAAGCCCGCCCATTTGCCAGGCGACGCC 213  
 QY 75 ---LysProLysAlaLysGluSerAsnThrLysGlnMetAsnLeuPheGlnArgAla 93  
 DB 214 CGCGCCCGCGAAGCGGCGGCGGCAAGAGCAGCTCAACTGTTCCAGCGCGCGCG 273  
 QY 94 AlaAlaLeuAspAla---AlaGluGlyPheLeuValSer---HisGluLysLeuHis 111  
 DB 274 CGCGCCCGCGTCAGCGCTTCAGGAAGGGTTCTGGCCCAACGCTCTCGAGGCGCCAC 333  
 QY 112 ProLeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValaen 131  
 DB 334 GGGCTGCCACAGCGCGCGACCGCGCTGCAGATCGCGGCAACTTCGCGCGCTCGGG 393  
 QY 132 GluGlnProValArgArgAsnLeuProValValGlyLysLeuProAspSerIleLysGly 151  
 DB 394 GAGAGCGCGCGGTGCACGAGCTCCCGCTCTCCGCGCGCATCCCGCCCTTCATCAGCGG 453  
 QY 152 ValTyrValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHisHisPhePhe 171  
 DB 454 GTCTACGCGCGCAACGGCGCCCAACCCCTGCTTCGACCCCGTCCGCGGCGCACCATCTTC 513  
 QY 172 AspGlyAspGlyMetValHisAlaValLysPheGluHisGlySerAla---SerTyrAla 190  
 DB 514 GACGCGACGCGATGTCACGCGCTCGGATACGCAACGCGCGCGCGAGTCTTCACGCC 573  
 QY 191 CysArgPheThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPhe 210  
 DB 574 TGCCGCTTCAGGAGACCGCGCGCTGCGCAGGAGCGCGGATCGCGCGCGCTCTTC 633  
 QY 211 ProLysAlaIleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyr 230  
 DB 634 CCCAAGGCCATTGGCGAGCTGCACGGGCACTCCGGGATCGCGGCTCGCCCTGTCTTAC 693  
 QY 231 AlaArgAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsnAlaGly 250

```
Db 694 GCGCGCGCGCTGGCGCTCGTGGACCCCTCGCGCGGCACCGCGTGGCCAAACGCGGC 753
Qy 251 LeuValTyrPheAsnGlyArgLeuLeuAlaMetSerGluAspLeuProTyrGlnVal 270
Db 754 CTGCTACTTCAACGGCGCGCTGCTGCCATGTCCGAGGACGACCTCCCTCCACAGCTC 813
Qy 271 GlnIleThrProAsnGlyAspLeuThrValGlyArgPheAspPheAspGlyGlnLeu 290
Db 814 CGGTGGCGGACGCGGACCTCGAGACCTCGCGCTCGCGCTCGAGCTTCGACGGGAGCTC 873
Qy 291 GluSerThrMetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeu 310
Db 874 GGCTGGCGCATGATCGCGCACCCCAAGCTGGACCGGCCACCGGGAGCTCCACGCGCTC 933
Qy 311 SerTyrAspValSerLysProTyrIleuLysTyrPheArgPheSerProAspGlyThr 330
Db 934 AGTACGACGTCATCAGAGCGCGTACTCCTCAAGTACTTCTACTTCAGCGCCGACGCGACC 993
Qy 331 LysSerProAspValGluIleGlnLeuAspGlnProThrMetMetHisAspPheAlaIle 350
Db 994 AAGTCGACGACGTGGAGATCCGCTGGAGCAGCCCGACGATGATCCAGACTTCGCCATC 1053
Qy 351 ThrGluAsnPheValValProAspGlnGlnValValPheLysLeuProGluMetIle 370
Db 1054 ACCGAGAACTTCGTGGTGTGCCCGCACCCAGGTGGTGTTCAGACTCCAGGAGTGTCTG 1113
Qy 371 ArgGlyGlySerProValValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAsp 390
Db 1114 CGCGCGCGTCCCGCTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1173
Qy 391 LysTyrAlaGluAspSerSerAsnIleLysTrpIleAspAlaProAspCysPheCysPhe 410
Db 1174 AAGCAGCGCGGACGCGCTGGAGATGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1233
Qy 411 HisLeuTrpAsnAlaTrpGluGluProGluThrAspGluValValIleGlySerCys 430
Db 1234 CACCTGTGAACGGTGGGAGACGAGCGCGGCGGAGGTGGTGGTGGTGGTGGTGGTGGTGG 1293
Qy 431 MetThrProProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSer 450
Db 1294 ATGACCCCGCGGACTCCTCAACGAGTCCGACGAGCGCTGGAGAGCGCTGGAGAGCGTGC 1353
Qy 451 GluIleArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGlu 470
Db 1354 GAGATCCCGCTGGACCGCGCGCGCGCTGCCAGCGCGCGCGCTGCTGCCG---CCG 1410
Qy 471 AspGlnGlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThr 490
Db 1411 TCGCAGCAGGAGAACCTCGAGGTGGCATGTGTGAACCGCAACCTGCTGGGCGCGAGAGC 1470
Qy 491 LysPheAlaTyrLeuAlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysVal 510
Db 1471 CGGTACCGGTACCTCGCGTGGGAGCGGTGGCCCAAGAGAGTGGGGTTCGCCCAAGGAG 1530
Qy 511 AspLeuThrThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGlu 530
Db 1531 GACCTGTCCACGGGGAGCTCACCAGTTCGAGTACGGCGGCGCGGTTCGGCGGGCGAG 1590
Qy 531 ProLeuPheLeuProGlyGluGlyGlyGlu-----GluAspGluGlyTyr 545
Db 1591 CCCTGCTTCGTTCCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGTAC 1650
Qy 546 IleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAla 565
Db 1651 GTGCTCACTTCGTCCACGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGT 1710
Qy 566 ValSerLeuGluValGluAlaThrValLysLeuProSerArgValProTyrGlyPheHis 585
Db 1711 GCCGACATCCCGCTGGAGCCACGGTTCAGCTCCGCTCCCGCTCCCGCTTCGCTTCAC 1770
Qy 586 GlyThrPheIleGlyAlaAspLeuAlaLysGln 597
Db 1771 GGCACCTTCATCACGGCGCAGAGCTCGAGGCCAC 1806
```

## RESULT 6

AAD09398 standard; cDNA; 1734 BP.

AAD09398;

10-SEP-2001 (first entry)

Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED5 cDNA.

Neoxanthin cleavage enzyme; AtNCED5; abscisic acid; ABA; herbicide;  
stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
plant growth protectant; ss.

Arabidopsis thaliana.

Key Location/Qualifiers

CDS 1..1734

/\*tag= a

/product= "Arabidopsis thaliana AtNCED5 protein"

EP1116794-A2.

18-JUL-2001.

11-JAN-2001; 2001EP-0300218.

13-JAN-2000; 2000JP-0010056.

11-JAN-2001; 2001JP-0003476.

(RIKE ) RIKEN KK.

Iuchi S, Kobayashi M, Shinozaki K;

WPI; 2001-400081/43.

P-PSDB; AAE04786.

A DNA encoding a protein with a neoxanthin cleavage activity for  
producing transgenic plants with improved or decreased stress tolerance

Claim 3; Page 46-49; 101pp; English.

The invention relates to neoxanthin cleavage enzymes and their  
corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key  
role in endogenous abscisic acid (ABA) biosynthesis under drought stress.  
Neoxanthin cleavage enzyme is used for improving stress tolerance in a  
plant when expressed in a plant cell. The invention also relates to  
methods for increasing or decreasing stress tolerance in a plant by  
introducing the DNA into the plant, and a transgenic plant into which a  
neoxanthin cleavage enzyme is introduced. The improvement of stress  
tolerance in plants is useful, for example in plant breeding. Neoxanthin  
cleavage enzyme genes are useful for producing transgenic plants. An arid  
land can be improved by growing transformant weed for several years and  
then removing the weed by specifically lowering stress tolerance in the  
weed by inducing an inducible promoter. The present cDNA sequence encodes  
Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED5 protein.  
The AtNCED5 cDNA is obtained from an Arabidopsis plant-derived cDNA  
library using a cDNA of the CPRD65 (CowPea Responsive to Dehydration)  
gene isolated from cowpea plant as a probe.

SQ Sequence 1734 BP; 464 A; 390 C; 441 G; 439 T; 0 other;

Alignment Scores:

Pred. No.:	1..2e-152	Length:	1734
Score:	1663.50	Matches:	323
Percent Similarity:	72.56%	Conservative:	100
Best Local Similarity:	55.40%	Mismatches:	137
Query Match:	52.81%	Indels:	23
DB:	22	Gaps:	9

US-09-758-269-6 (1-599) x AAD09398 (1-1734)





PR 23-MAR-1999; 99US-0135788;  
PR 25-MAR-1999; 99US-0136264;  
PR 29-MAR-1999; 99US-0136785;  
PR 01-APR-1999; 99US-0127462;  
PR 06-APR-1999; 99US-0128234;  
PR 08-APR-1999; 99US-0128714;  
PR 16-APR-1999; 99US-0129845;  
PR 19-APR-1999; 99US-0130077;  
PR 21-APR-1999; 99US-0130444;  
PR 23-APR-1999; 99US-0130510;  
PR 23-APR-1999; 99US-0130891;  
PR 28-APR-1999; 99US-0131449;  
PR 30-APR-1999; 99US-0132048;  
PR 30-APR-1999; 99US-0132407;  
PR 04-MAY-1999; 99US-0132484;  
PR 05-MAY-1999; 99US-0132485;  
PR 06-MAY-1999; 99US-0132486;  
PR 06-MAY-1999; 99US-0132487;  
PR 07-MAY-1999; 99US-0132863;  
PR 11-MAY-1999; 99US-0134256;  
PR 14-MAY-1999; 99US-0134218;  
PR 14-MAY-1999; 99US-0134219;  
PR 14-MAY-1999; 99US-0134221;  
PR 14-MAY-1999; 99US-0134370;  
PR 18-MAY-1999; 99US-0134768;  
PR 19-MAY-1999; 99US-0134941;  
PR 20-MAY-1999; 99US-0135124;  
PR 21-MAY-1999; 99US-0135353;  
PR 24-MAY-1999; 99US-0135622;  
PR 25-MAY-1999; 99US-0136021;  
PR 27-MAY-1999; 99US-0136392;  
PR 28-MAY-1999; 99US-0136782;  
PR 01-JUN-1999; 99US-0137222;  
PR 03-JUN-1999; 99US-0137528;  
PR 04-JUN-1999; 99US-0137503;  
PR 07-JUN-1999; 99US-0137724;  
PR 08-JUN-1999; 99US-0138094;  
PR 10-JUN-1999; 99US-0138540;  
PR 10-JUN-1999; 99US-0138847;  
PR 14-JUN-1999; 99US-0139119;  
PR 16-JUN-1999; 99US-0139453;  
PR 16-JUN-1999; 99US-0139453;  
PR 17-JUN-1999; 99US-0139492;  
PR 18-JUN-1999; 99US-0139454;  
PR 18-JUN-1999; 99US-0139455;  
PR 18-JUN-1999; 99US-0139456;  
PR 18-JUN-1999; 99US-0139457;  
PR 18-JUN-1999; 99US-0139458;  
PR 18-JUN-1999; 99US-0139459;  
PR 18-JUN-1999; 99US-0139460;  
PR 18-JUN-1999; 99US-0139461;  
PR 18-JUN-1999; 99US-0139462;  
PR 18-JUN-1999; 99US-0139463;  
PR 18-JUN-1999; 99US-0139463;  
PR 18-JUN-1999; 99US-0139750;  
PR 18-JUN-1999; 99US-0139763;  
PR 21-JUN-1999; 99US-0139817;  
PR 22-JUN-1999; 99US-0139889;  
PR 23-JUN-1999; 99US-0140353;  
PR 23-JUN-1999; 99US-0140353;  
PR 24-JUN-1999; 99US-0140695;  
PR 28-JUN-1999; 99US-0140823;  
PR 29-JUN-1999; 99US-0140991;  
PR 30-JUN-1999; 99US-0141287;  
PR 01-JUL-1999; 99US-0141842;  
PR 01-JUL-1999; 99US-0142154;  
PR 02-JUL-1999; 99US-0142055;  
PR 06-JUL-1999; 99US-0143290;  
PR 08-JUL-1999; 99US-0142803;  
PR 09-JUL-1999; 99US-0142920;  
PR 12-JUL-1999; 99US-0142977;  
PR 13-JUL-1999; 99US-0143544;  
PR 14-JUL-1999; 99US-0143624;  
PR 15-JUL-1999; 99US-0144005;  
PR 16-JUL-1999; 99US-0144085;  
PR 16-JUL-1999; 99US-0144086;  
PR 19-JUL-1999; 99US-0144325;  
PR 19-JUL-1999; 99US-0144331;  
PR 19-JUL-1999; 99US-0144332;  
PR 19-JUL-1999; 99US-0144333;  
PR 19-JUL-1999; 99US-0144334;  
PR 19-JUL-1999; 99US-0144335;  
PR 20-JUL-1999; 99US-0144352;  
PR 20-JUL-1999; 99US-0144632;  
PR 20-JUL-1999; 99US-0144884;  
PR 21-JUL-1999; 99US-0144814;  
PR 21-JUL-1999; 99US-0145086;  
PR 21-JUL-1999; 99US-0145088;  
PR 22-JUL-1999; 99US-0145085;  
PR 22-JUL-1999; 99US-0145087;  
PR 22-JUL-1999; 99US-0145089;  
PR 22-JUL-1999; 99US-0145192;  
PR 23-JUL-1999; 99US-0145145;  
PR 23-JUL-1999; 99US-0145218;  
PR 23-JUL-1999; 99US-0145224;  
PR 26-JUL-1999; 99US-0145276;  
PR 27-JUL-1999; 99US-0145913;  
PR 27-JUL-1999; 99US-0145918;  
PR 27-JUL-1999; 99US-0145919;  
PR 28-JUL-1999; 99US-0145951;  
PR 02-AUG-1999; 99US-0146386;  
PR 02-AUG-1999; 99US-0146388;  
PR 03-AUG-1999; 99US-0146389;  
PR 03-AUG-1999; 99US-0147038;  
PR 04-AUG-1999; 99US-0147204;  
PR 04-AUG-1999; 99US-0147302;  
PR 05-AUG-1999; 99US-0147192;  
PR 05-AUG-1999; 99US-0147260;  
PR 06-AUG-1999; 99US-0147303;  
PR 06-AUG-1999; 99US-0147416;  
PR 09-AUG-1999; 99US-0147493;  
PR 09-AUG-1999; 99US-0147935;  
PR 10-AUG-1999; 99US-0148171;  
PR 11-AUG-1999; 99US-0148319;  
PR 12-AUG-1999; 99US-0148341;  
PR 13-AUG-1999; 99US-0148565;  
PR 13-AUG-1999; 99US-0148684;  
PR 16-AUG-1999; 99US-0149368;  
PR 17-AUG-1999; 99US-0149175;  
PR 18-AUG-1999; 99US-0149426;  
PR 20-AUG-1999; 99US-0149722;  
PR 20-AUG-1999; 99US-0149723;  
PR 20-AUG-1999; 99US-0149929;  
PR 23-AUG-1999; 99US-0149902;  
PR 23-AUG-1999; 99US-0149930;  
PR 25-AUG-1999; 99US-0150566;  
PR 26-AUG-1999; 99US-0150884;  
PR 27-AUG-1999; 99US-0151065;  
PR 27-AUG-1999; 99US-0151066;  
PR 27-AUG-1999; 99US-0151080;  
PR 30-AUG-1999; 99US-0151303;  
PR 31-AUG-1999; 99US-0151438;  
PR 01-SEP-1999; 99US-0151930;  
PR 07-SEP-1999; 99US-0152363;  
PR 10-SEP-1999; 99US-0153070;  
PR 13-SEP-1999; 99US-0153758;  
PR 15-SEP-1999; 99US-0154018;  
PR 16-SEP-1999; 99US-0154039;  
PR 20-SEP-1999; 99US-0154779;  
PR 22-SEP-1999; 99US-0155139;  
PR 23-SEP-1999; 99US-0155486;  
PR 24-SEP-1999; 99US-0155659;  
PR 28-SEP-1999; 99US-0156458;  
PR 29-SEP-1999; 99US-0156596;  
PR 04-OCT-1999; 99US-0157117;  
PR 05-OCT-1999; 99US-0157753;  
PR 06-OCT-1999; 99US-0157865;









PS Claim 1; Page 94-97; 135pp; English.

CC This invention relates to defence-related signalling genes isolated from  
CC the sunflower (*Helianthus annuus*). The genes encode a neoxanthin cleavage  
CC enzyme (NCE), an amino acid permease (AAP) and a glutamic acid rich  
CC protein (GRP). The signalling gene is useful for increasing the  
CC resistance of a plant to a pathogen such as fungus, virus, bacterium,  
CC nematode or insect (e.g. European corn borer), preferably  
CC *Sclerotinia* spp., *Phoma* spp., or *Phomopsis* spp., by stably incorporating a  
CC construct containing the gene into the genome of the plant. The gene is  
CC useful for regulating gene expression in a plant, in response to a  
CC stimulus such as infection with a pathogen, damage from a pathogen,  
CC hydrogen peroxide, jasmonic acid, methyl jasmonate, salicylic acid,  
CC oxalic acid or expression of a gene encoding oxalic acid oxidase. The  
CC genes are also useful for stem-preferred regulation of gene expression in  
CC a plant. The genes are useful in agriculture, particularly in the  
CC breeding of crop plants with improved agronomic traits, for modifying  
CC abscisic acid (ABA) metabolism and for modifying amino acid transport and  
CC content in plants. The present sequence represents cDNA encoding the  
CC sunflower neoxanthin cleavage enzyme (NCE).

XX SQ Sequence 1950 BP; 535 A; 456 C; 453 G; 506 T; 0 other;

#### Alignment Scores:

Pred. No.:	9.37e-84	Length:	1950
Score:	960.50	Matches:	209
Percent Similarity:	53.9%	Conservative:	109
Best Local Similarity:	35.4%	Mismatches:	214
Query Match:	30.4%	Indels:	57
DB:	22	Gaps:	13

US-09-758-269-6 (1-599) x AAF77206 (1-1950)

QY 51 SerSerAlaLeuHisThrProProAlaLeuHisPheProLysGlnSerSerAsnSerPro 70  
DB 49 TCAACTGTTCCACCCACCACCA-----TCATCGTCTCACCG 87  
QY 71 AlaIleValVal-----LysProLysAlaLysGlnSer 81  
DB 88 CCATCTCGCGCTAAGAGTCTTTTTCAGTCAGGACTGAAGAAACACAGAAACCGTCCACCACT 147  
QY 82 -----AsnThrLysGln-----MetAsnLeuPheGln 90  
DB 148 ACCGCCACCCAAAGAGCCTAGTCGGACGCGAAACCAACCCCTGAATATAGANAA 207  
QY 91 ArgAlaAlaAlaAlaLeuAspAla-----AlaGluGly 102  
DB 208 CGTGAAGCTTTCAGTCGGTGGTGCAGTCTTTACCGGCCACCCTTTAATGCTTTTCAT 267  
QY 103 PheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerValGln 122  
DB 268 AGCATCATTAATAACTTTATTGATCCACCGCGGAGGGTCTCAGTTGATCCAAACACCGTT 327  
QY 123 IleAlaGlyAsnPheAlaProValAsnGlnGlnProValArgAsnLeuProValVal 142  
DB 328 TTGCTGATAATTTTTCACCGGTGACAACTCCCTCCGACTGCTGACTGCAAGTCATCGAG 387  
QY 143 GlyLysLeuProAspSerIleLysGlyValValArgAsnGlyAlaAsnProLeuHis 162  
DB 388 GGCACATGCCAAGTTCCTGACCGTCTTACTTCCGTAATAGTCCGAACCCGCAATTC 447  
QY 163 GluProValThrGlyHisPheAspGlyAspGlyMetValHisAlaValLysPhe 182  
DB 448 CTTCCGCGAGGACCCCTACCACTCTTCGATGSCGATGCGTCTCATGCTATTCTGATC 507  
QY 193 GluHisGlySerAlaSerTyAlaCysArgPheThrGlnThrAsnArgPheValGlnGlu 202  
DB 508 TCCATGAAAGCTTCGTTATGTAGCCGATCATCAATCAAAATATTTCATATAGAG 567  
QY 203 ArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHisThr--- 221  
DB 568 AAGAAAGCGGGATTCCTCCATTTATCCAAACGTGTTTTCAGGGTTTAAATGGTGTGACTGCC 627

QY 222 GlyIleAlaArgLeuMetLeuPheTyAlaArgAlaAlaAlaGlyIleValAspProAla 241  
DB 628 TCTGCAGCTCGATCGCAGTCACTCCGCGCGATTTTGGCTGGACAATTTTACCCCA 687  
QY 242 HisGlyThrGlyValAlaAsnAlaGlyLeuValTyPheAsnGlyArgLeuLeuAlaMet 261  
DB 688 AAGGTATTGGTCTAGCCCAATACCACTGCGCTTTTGGCAACAGACTTTTGTCTCG 747  
QY 262 SerGluAspAspLeuProTyArgGlnValGlnIleThrProAsnGlyAspLeuLysThrVal 281  
DB 748 GGAGAGTCGGATCTCCATATCCCGTCAACTAGCGCCGCGGTCGATAGTACCCGTC 807  
QY 282 GlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysValAsp 301  
DB 808 GGACGTGAGGACTTCGACGGCAAACTATTCATGACATGACCGTCACCCAAAATTCGAT 867  
QY 302 ProGluSerGlyGluLeuPheAlaLeuSerTyArgPheValValSerLysProTyLeuLys 321  
DB 868 CAGTAAACGAAAGAGCTTTTGGCTTTCGTTACGGTCCAGTC---CCCCCTTTCTTAACC 924  
QY 322 TyrPheArgPheSerProAspGlyThrLysSerProAspValGluIle---GlnLeuAsp 340  
DB 925 TTTTTCGTTTCAACGAAACCGAGAAAAACAAGCCGATGTCCCGATCTTCTCATATGACA 984  
QY 341 GlnProThrMetHisAspPheAlaIleThrGluAsnPheValValProAspGln 360  
DB 985 AGCCCGTCTGTTCTCCACGACTTCGCCATCACCAAAAACTACCGCATTTTCCCGAGATC 1044  
QY 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyAspLys 380  
DB 1045 CAAATCGGGATGAGCCCAATCGAGATGCTGGTGGGGATCCCCGGTTAGCGCGACGCT 1104  
QY 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyAlaGluAspSerSerAsnIleLys 400  
DB 1105 GGAAAGGTGCTCGCTCGGTCGGTGTGATCCTCGGTACGCGAAGACGAGTCCGAGATGAAG 1164  
QY 401 TrpIleAspAlaProAspCysPheHisLeuTrpAsnAlaTrpGluGluProGlu 420  
DB 1165 TGGTTTTCAGGTTCGGGTTTAAATGTGATACATGTCATCAATGTCATGCGAGGAGGATGCC 1224  
QY 421 ThrAspGluValValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440  
DB 1225 GGAGATACGGTGGTGTGATGGTCCGAAACATATATTCGTTGAACATACATCGTTGGAAGA 1284  
QY 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460  
DB 1285 ATGGAT---TTGATTCATGATCGGTTGAGAAAGTGAAGATTAATCTGAAACACAGAAATG 1341  
QY 461 SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480  
DB 1342 GTATCCCGGACCCGCTTCAACCCGG-----AATCTTGACTTTTCGGT 1386  
QY 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyLeuAlaLeuAlaGluPro 500  
DB 1387 TTAATCCGGCTTTGTTCCGCTTAAACACAGGTATATGTATGTTGGAGTTGGTATCGG 1446  
QY 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGlu----- 516  
DB 1447 ATGCCAAGATCTCTGGTGTGCTCAAGCTAGATGTGTCATCTCCGAAGCAGACCGTCGC 1506  
QY 517 -----ValLysLysHisLeuTyArgLysAsnArgTyGlyGlyGluProLeuPhe 533  
DB 1507 GAATGCATAGTTGCTAGCCGATGTTGGGCTGTTGTTACGTTGGTGGTGAACCATCTTT 1566  
QY 534 LeuProGlyGlu-----GlyGlyGluAspGluGlyTyIleLeuCysPhe 549  
DB 1567 GTAGCTAGGAGCCAGACCAACCCCGATGCGATGAGGATGATGGGTATGTTGTTCTTAT 1626  
QY 550 ValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaValSer----- 567  
DB 1627 GTGATACGAGAACACCGGTGAGTCAAGATTCGTTGGTGGAGCGCTAAGTACCGAGG 1686  
QY 568 LeuGluValGluAlaThrValLysLeuProSerArgValProTyGlyPheHisGlyThr 587

```
Db 1687 CTTGAGATCGTCGATCGGTGAAGCTGCCCCACCGGGTACCATATGTTTTCATGGGCTC 1746
Oy 588 PheIleGlyAlaAspAspLeuAlaIys 596
Db 1747 TTTGTTAGAGAAAGTGACGTTTAACAAG 1773

RESULT 11
AAC36083
ID AAC36083 standard; DNA; 1777 BP.
AC AAC36083;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 12487.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
```







```
Db 439 GGATTGCTAATGCTCAATGCTCAACAGCTGAGAACGAACTCAAAATATTGGACAACT 498
Qy 242 HisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuAlaMet 261
Db 499 TATGAAATGGAACCTGCAATACAGCACTCGTATATACCATGGAACCTCTAGCAATTA 558
Qy 262 SerGluAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThrVal 281
Db 559 CAGGAGGCGCATAGCGTACGTCATCAAAAGTTTGGAAAGATGAGACCTGCAAACTCTT 618
Qy 282 GlyArgPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysValAsp 301
Db 619 GGTATATAGATTATGACAAGATTGACCCACTCTCTTCACTGCTCACCACAAAGTTGAC 678
Qy 302 ProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeuLys 321
Db 679 CCGGTTACGGGTGAATGTTTACATTCGGCTATTTCG---CATACGCCACTTATCTCACA 735
Qy 322 TyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAspGln 341
Db 736 TACAGAGTTATCTCGAAAGATGCAATTATGCAATGCCAGTCCCAATTACTATATCAGAG 795
Qy 342 ProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGlnGln 361
Db 796 CTTATCATGATCATGATTTTCTATTTACTGAGCTTATGCAATCTTCAATGATCTTCTCT 855
Qy 362 ValValPheLeuProGluMetIleArgGlyGlySerProValVal---TyrAspLys 380
Db 856 ATGCACCTTCAGGCCAAGCAATGTTGAAGAGAGAAATGATATACTACTCATTTGATCCC 915
Qy 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnLys 400
Db 916 ACAAAAAAGCGCTCGTTTGGTCTCTTCCACGCTATGCCAAGATGCAACTTATGATTAGA 975
Qy 401 TrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluProGlu 420
Db 976 TGGTTTGAGCTTCCCAACTGCTTTATTTTCCAAACGCGCAATGCTTGGGAA-----GAA 1029
Qy 421 ThrAspGluValValIleGlySerCysMetThrProProAsp----- 435
Db 1030 GAGGATGAAGTCTCTCATCACTTGTCTGCTTGAGATCGAGATCTTGACATGTCAGT 1089
Qy 436 SerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsn 455
Db 1090 GGGAAAGTGAAGAAAAAATCTGAAATTTTGGCAACGAACTGTACGAAATGAGATTCAAC 1149
Qy 456 LeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnValAsn 475
Db 1150 ATGAAAAAGGGCTCAGCTTCTCAAAAAAACTATCCGCATCTCGC----- 1194
Qy 476 LeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeu 495
Db 1195 GTTGATTTCCTCCAGATCATGAGTGTACACCGGAAGAAACAGAGATACGTATATGA 1254
Qy 496 AlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAsp----- 511
Db 1255 ACAATTCTGGACAGTATCGAAAGGTACCGGAATCATCAAGTTTGATCTGCATCGAGAA 1314
Qy 512 -----LeuThrThrGlyGluValLysLysHisLeuTyr----- 522
Db 1315 CTGAGACAGGGAAGAAATGCTGGAAGTAGAGGTAATATCAAGGAATATATACACCTG 1374
Qy 523 GlyAspAsnArgTyrGlyGlyGluProLeuPheLeuProGlyGlyGlyGluGluAsp 542
Db 1375 GGAGAGGAGGAGATATGTTTCAGAGCTATCTATGTTCCGCTGAGACAGCAGAAAGAC 1434
Qy 543 GluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIle 562
Db 1435 GACGGTTACTTGTATTTCTTTCTTCTCATGATGAAACACAGGGAATCATCGTGTGCTGTG 1494
Qy 563 ValAsnAlaValSerLeuGluValGlu-----AlaThrValLysLeuProSerArgVal 580
```

```
Db 1495 ATAGACGCCAAAAACAATGTCCGCTGAACCCGGTGGCAGTGGTGGAGCTGCCGACAGGTC 1554
Qy 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGlnValVal 599
Db 1555 CCATATGCTTCCATGCTGTTGTTGTACAGAGGAACAACTCCAGGAACAAACTCTT 1611

RESULT 13
AAD09402
ID AAD09402 standard; cDNA; 1617 BP.
XX
AC AAD09402;
XX
DT 10-SEP-2001 (first entry)
XX
DE Arabidopsis thaliana neoxanthin cleavage enzyme cDNA.
XX
KW Neoxanthin cleavage enzyme; abscisic acid; ABA; herbicide;
KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;
KW plant growth protectant; ss.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT CDS 1..1617
FT /*tag= a
FT /product= "Arabidopsis neoxanthin cleavage enzyme"
XX
PN EP1116794-A2.
PD 18-JUL-2001.
XX
PF 11-JAN-2001; 2001EP-0300218.
XX
PR 13-JAN-2000; 2000JP-0010056.
PR 11-JAN-2001; 2001JP-0003476.
XX
PA (RIKE ) RIKEN KK.
PI Iuchi S, Kobayashi M, Shinozaki K;
XX
WPI: 2001-400081/43.
DR P-PSDB; AAE04790.
XX
A DNA encoding a protein with a neoxanthin cleavage activity for
producing transgenic plants with improved or decreased stress tolerance
-
PS Disclosure; Page 74-77; 101pp; English.
XX
CC The invention relates to neoxanthin cleavage enzymes and their
corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key
role in endogenous abscisic acid (ABA) biosynthesis under drought stress.
CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a
plant when expressed in a plant cell. The invention also relates to
methods for increasing or decreasing stress tolerance in a plant by
introducing the DNA into the plant, and a transgenic plant into which a
neoxanthin cleavage enzyme is introduced. The improvement of stress
tolerance in plants is useful, for example in plant breeding. Neoxanthin
cleavage enzyme genes are useful for producing transgenic plants. An arid
land can be improved by growing transformant weed for several years and
then removing the weed by specifically lowering stress tolerance in the
weed by inducing an inducible promoter. The present cDNA sequence encodes
Arabidopsis thaliana neoxanthin cleavage enzyme related to the invention.
XX
SQ Sequence 1617 BP; 480 A; 348 C; 368 G; 421 T; 0 other;

Alignment Scores:
Pred. No.: 1,426-81 Length: 1617
Score: 937.00 Matches: 207
Percent Similarity: 55.81% Conservative: 105
Best Local Similarity: 37.03% Mismatches: 199
Query Match: 29.75% Indels: 48
DB: 22 Gaps: 13
```



XX WPI; 2000-579369/54.  
XX New isolated polynucleotide encoding a plant transcription factor for  
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
PT having modified gene expression or modified activity of a polypeptide  
PT  
XX  
XX  
XX Claim 1; Page 490; 747pp; English.  
XX The present invention relates to novel plant transcription factors from  
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
CC sequence for one such transcription factor. The transcription factor may  
CC be used to produce a plant having modified gene expression such as a  
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
CC mahogany species or to modify the activity of a polypeptide in a plant.  
CC The transcription factors of the present invention are members from the  
CC following families of regulatory proteins: bZIP, bZIP family of G-box  
CC binding factors, basic helix-loop-helix zipper,  
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
CC and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
CC and MYB.  
XX  
SQ Sequence 443 BP; 71 A; 177 C; 117 G; 77 T; 1 other;  
Alignment Scores:  
Pred. No.: 3.38e-50 Length: 443  
Score: 608.00 Matches: 115  
Percent Similarity: 87.76% Conservative: 14  
Best Local Similarity: 78.23% Mismatches: 18  
Query Match: 19.30% Indels: 0  
DB: 21 Gaps: 0  
US-09-758-269-6 (1-599) x AAC56678 (1-443)  
QY 208 ProValPheProLysAlaIleGlyLeuHisGlyHisThrGlyLeuAlaArgLeuMet 227  
DB 3 CCGGCTTTCCNCAAGCCATCGCGAGCTCCACGCCCACTCGGCGATCGCGGCTCATG 62  
QY 228 LeuPheTyrAlaArgAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAla 247  
DB 63 CTCTTCTACGCCCGAGGCTCTCGGCTCTGTCGACCCGAGATGGCGTTCGCG 122  
QY 248 AsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAlaMetSerGluAspLeuPro 267  
DB 123 AACGGCGGCTCGTGACTTCGACGGCCACCTCTCTCGGATGTCGAGAGACGACCTCCC 182  
QY 268 TyrGlnValGlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAspLeu 287  
DB 183 TACCACGTGCGGTCACGCGCTTCGCGGACCTCGAGACCGTCCGCGCTACGACTTCGCC 242  
QY 288 GlyGlnLeuGluSerThrMetIleAlaHisProLysValAspProGluSerGlyGluLeu 307  
DB 243 GGCCAGCTCGACTCTCCGATGATGCCACCCGAGATCGACCCGGCTTCGCGGAGATG 302  
QY 308 PheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerPro 327  
DB 303 TTCGCCCTCAGCTACGACGTGTCGCGAAGCGGTACTCTCACTTCCGATTCTCCAAG 362  
QY 328 AspGlyThrLysSerProAspValGluIleGlnLeuAspGlnProThrMetMetHisAsp 347  
DB 363 GACGCGGAGAAGTCCCGGAGCTCGAGATCCCTGCTGAGTCCCTGGCTGAGCCCATGATGCACGAT 422  
QY 348 PheAlaIleThrGluAsnPhe 354  
DB 423 TTCGCCATACCGAGCGCTTT 443  
RESULT 15  
ID AAC82706/c  
XX AAC82706 standard; DNA; 492 BP.  
XX  
AC AAC82706;  
XX

DT 15-MAR-2001 (first entry)  
XX  
DE Rice abscisic acid synthesis associated DNA SEQ ID NO 5.  
XX  
KW Rice; abscisic acid synthesis; VP14-like gene; drought resistance; ds.  
XX  
OS Oryza sativa.  
XX  
PN WO200071727-A1.  
XX  
PD 30-NOV-2000.  
XX  
PF 25-MAY-1999; 99WO-JP02734.  
XX  
PR 25-MAY-1999; 99WO-JP02734.  
XX  
PA (NORO) JAPAN MIN AGRIC FORESTRY & FISHERIES.  
PA (BIOO-) BIO-ORIENTED TECHNOLOGY RES ADVANCEMENT.  
XX  
PI Hirochika H, Sakamoto K;  
XX  
DR WPI; 2001-032042/04.  
XX  
PT Oligonucleotide encoding gene for regulating abscisic acid synthesis in  
PT plants, useful for constructing e.g. genetically-modified rice with  
PT drought resistance and ear-germination resistance  
XX  
PS Disclosure; Page 50; 55pp; Japanese.  
XX  
CC This invention describes a novel polynucleotide sequence (I) which  
CC encodes a protein capable of regulating the synthesis of abscisic acid.  
CC The invention also describes (1) an oligonucleotide encoding a protein  
CC of a gene obtained by controlling the expression of a VP14-like gene;  
CC (2) a vector containing the oligonucleotide ligated operably to the  
CC regulation sequence; (3) a plant transformed with the vector; and  
CC (4) a method for regulating abscisic acid synthesis in a plant including  
CC the transfer of the above oligonucleotide to it. The gene is useful for  
CC constructing drought resistant rice.  
SQ Sequence 492 BP; 90 A; 126 C; 186 G; 88 T; 2 other;  
Alignment Scores:  
Pred. No.: 1.45e-48 Length: 492  
Score: 592.00 Matches: 115  
Percent Similarity: 90.28% Conservative: 15  
Best Local Similarity: 79.86% Mismatches: 13  
Query Match: 18.79% Indels: 2  
DB: 22 Gaps: 0  
US-09-758-269-6 (1-599) x AAC82706 (1-492)  
QY 210 PheProLysAlaIleGlyLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPh 229  
DB 434 TTTCTTAAGGCGATAGGTGAGCTTCATGGCCACTCCGGGCGATCGCGCCCTTGCTCTGT 375  
QY 229 eTyrAlaArgAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsnAl 249  
DB 374 CTACGGCGCGCGGCTTC-GGCTCTCTCGACCGCTCACACGGCAGCGGCTCGCCNAGCG 316  
QY 249 aGlyLeuValTyrPheAsnGlyArgLeuLeuAlaMetSerGluAspLeuProTyrGl 269  
DB 315 CGGCTCATCTACTTCAACGGCAGGCTCTCGCATGTGCGAGGACGACCTCCCTACCA 256  
QY 269 nValGlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAspLeuGly 289  
DB 255 GGTGCGGCTGACCGCGGCGGCTTCGAGACCGTCCGCGGCTCGACGCTTCGACGGCA 196  
QY 289 nLeuGluSerThrMetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAl 309  
DB 195 GCTGGCTGCGCCATGATCGCGGACCCCAAGCTCGACCCGCGCCACCGGAGAGCTCCAGCG 136  
QY 309 aLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGl 329  
|||||

Db	135	GCTCAGCTACGACGTGATCAAGAAGCCGTACCTCAAGTACTTCTACTTCGGCCCGACGG	76
Qy	329	YThriYsSerProAspValGluIleGlnLeuAspGlnProThrMetMetHisAspPheAl	349
Db	75	CACCAAGTCGGCCGCGACGTCGAGATCCCGCTCGACGAGCCACCATGATCCACGACTTCCG	16
Qy	349	aileThrGlu	352
Db	15	AATTACTGAG	6

Search completed: November 15, 2003, 15:07:18  
Job time : 522 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 15, 2003, 12:55:19 ; Search time 103 Seconds  
(without alignments)  
2566.879 Million cell updates/sec

Title: US-09-758-269-6  
Perfect score: 3150  
Sequence: 1 MASFTATAAASGRWLGNNHT.....VPYGHGTFIGADDLAKQV 599

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-O=/cn2\_1/uspto.spool/US09758269/runat.14112003.135757.28769/app.query.fasta.1.775  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOFCLO=0  
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09758269 @CN 1.1 56 @runat.14112003.135757.28769 -NCPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cn2\_6/ptodata/2/ina/PTUS\_COMB.seq:\*  
6: /cn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	494.5	15.7	4403765	3	US-09-103-840A-2
2	494.5	15.7	4411529	3	US-09-103-840A-1
3	475.5	15.1	4403765	3	US-09-103-840A-2
4	475.5	15.1	4411529	3	US-09-103-840A-1
5	419	13.3	32679	4	US-08-976-063E-1
6	411.5	13.1	1518	4	US-08-976-063E-21
7	211	6.7	2629	1	US-08-200-807-1
8	211	6.7	2629	1	US-08-488-305A-1
9	210	6.7	1724	3	US-09-385-259-1
10	210	6.7	1724	4	US-09-645-370-1
11	186	5.9	36063	4	US-08-311-731A-140
12	127.5	4.0	4242	4	US-09-252-991A-7056

13	127.5	4.0	10023	4	US-09-252-991A-6997	Sequence 6997, Ap
14	123.5	3.9	5163	3	US-08-700-651-1	Sequence 1, Appli
15	123.5	3.9	5163	3	US-08-928-361B-4	Sequence 4, Appli
16	123.5	3.9	5163	3	US-09-588-995A-4	Sequence 2, Appli
17	123.5	3.9	5318	3	US-08-700-651-2	Sequence 2, Appli
18	123.5	3.9	5318	3	US-08-928-361B-3	Sequence 3, Appli
19	123.5	3.9	5318	4	US-09-588-995A-3	Sequence 3, Appli
20	116.5	3.7	5511	3	US-08-928-361B-2	Sequence 2, Appli
21	116.5	3.7	5511	4	US-09-588-995A-2	Sequence 1, Appli
22	116.5	3.7	7334	3	US-08-928-361B-1	Sequence 1, Appli
23	116.5	3.7	7334	4	US-09-588-995A-1	Sequence 1, Appli
24	115	3.7	1389	4	US-09-252-991A-11721	Sequence 11721, A
25	115	3.7	2322	4	US-09-252-991A-11519	Sequence 11519, A
26	115	3.7	47981	4	US-09-679-279-1	Sequence 1, Appli
27	110	3.5	1635	4	US-09-252-991A-2493	Sequence 2493, Ap
28	106	3.4	1429	3	US-09-461-697-76	Sequence 76, Appl
29	106	3.4	1825	3	US-09-461-697-75	Sequence 75, Appl
30	106	3.4	1953	4	US-09-252-991A-759	Sequence 759, App
31	105.5	3.3	1416	4	US-09-107-532A-2051	Sequence 2051, Ap
32	105.5	3.3	4508	5	PCT-US93-06251-34	Sequence 34, Appl
33	105	3.3	1172	3	US-08-861-774E-17	Sequence 17, Appl
34	104.5	3.3	1389	4	US-09-328-352-3002	Sequence 3002, Ap
35	104	3.3	1341	4	US-09-350-756-2	Sequence 2, Appli
36	103	3.3	1479	4	US-09-252-991A-4180	Sequence 4180, Ap
37	103	3.3	1575	4	US-09-252-991A-2009	Sequence 2009, Ap
38	103	3.3	1581	4	US-09-252-991A-654	Sequence 654, App
39	103	3.3	3102	4	US-09-252-991A-4429	Sequence 4429, Ap
40	102.5	3.3	2607	2	US-08-907-166-1	Sequence 1, Appli
41	102.5	3.3	2607	4	US-09-391-340-1	Sequence 39, Appl
42	102.5	3.3	3257	4	US-09-585-173B-39	Sequence 111, App
43	102.5	3.3	14672	4	US-08-961-527-111	Sequence 12072, A
44	102	3.2	879	4	US-09-252-991A-12072	Sequence 14676, A
45	102	3.2	3396	4	US-09-252-991A-14676	

ALIGNMENTS

RESULT 1  
US-09-103-840A-2  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Alignment Scores:  
Pred. No.: 4,228-39 Length: 4403765  
Score: 494.50 Matches: 180  
Percent Similarity: 40.18% Conservative: 94  
Best Local Similarity: 26.33% Mismatches: 251  
Query Match: 15.70% Indels: 157  
DB: 3 Gaps: 27

US-09-758-269-6 (1-599) x US-09-103-840A-2 (1-4403765)

```
Qy 1 MetAlaSerPheThrAlaThr---AlaAlaValSerGly-----Arg 13
Db 751637 ATGGCCAGCAGGTGGCGACCGGATCGCGGTCTGCGGAGCGGTACGCGCGACATCGAAC 751696
Qy 14 TrpLeuGlyGlyAsnHisThrGlnProProLeuSerSerSerGlnSerSerAspLeuSer 33
Db 751697 TGGCGTACGCCCTTCTTCGGCTATGGCGCGATCAGTCGCGCATCCCGCCGAAATGGGTG 751756
Qy 34 TyrCysSer-----SerLeuProMetAlaSerArgValThrArgLysLeuAsnVal--- 50
Db 751757 TACACCGCATCGTCGAGGTCCCTGCTCGCGCGCCACCTTTCGGGGTCTGCAGCGCGTCG 751816
Qy 51 -----SerSerAlaLeuHisThrProAlaLeuHisPheProLysGlnSerSer 67
Db 751817 GCCCGGTGATCGTGGAGCAGTCGCGCGCGCGGTGC-----AGCAGCTCGTCTG 751864
Qy 68 AsnSerProAlaIleValVal-----LysProLysAlaLysGluSerAsn 82
Db 751865 CGAACACCGGCTCGGAGGTCTATCTTCCCATGTTCTCACCAAGGCGGTACCGTTCCTCAAT 751924
Qy 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAlaGluGly 102
Db 751925 ATCAGTGAATAACAATTTATAGGAGATCGGCATGACCAACCGCACAGCGCCGAA--- 751981
Qy 103 PheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerValGln 122
Db 751982 -----TCCCAAAACCCCATATCTCGAG 752002
Qy 123 IleAlaGlyAsnPheAlaProValAsnGlnGlnProValArgArgAsnLeuProValVal 142
Db 752003 -----GGCTCTTGGCGCGGTGAGCACCGAGTAACCTGCCACCGACCTGCCGTCACC 752056
Qy 143 GlyLysLeuProAspSerIleGlyGlyValTyrValArgAsnGlyAlaAsnProLeuHis 162
Db 752057 GCCCGCATTCGGAAACACCTACGACGGCGGTATCTGCTAAACGGCCCAACCGCGTCGCG 752116
Qy 163 Glu-----ProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaVal 180
Db 752117 GAGGTGACCCCGGCCACC---TACACTGGTTTACCGCGCGAGCGCATGGTGCACGGAGTC 752173
Qy 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200
Db 752174 CGGCTGCGCGACGGGAAGGCC-----CGCTGGTAT 752203
Qy 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGlu----- 216
Db 752204 CGCAATCGCTGGTCCGACACCGCGGTGTGCGCGCCCTTGGCGAGCCCATTTCCGCC 752263
Qy 217 ---LeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAla 235
Db 752264 CGGCTCACCCCGCGACCGGGATATCGAG----- 752293
Qy 236 GlyIleValAspProAlaHisGlyThrGlyValAlaAenAlaGlyLeuValTyrPheAsn 255
Db 752294 -----GGCGGTCCCAACACCAACGCTGCTGACCCACGCC 752326
Qy 256 GlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsn 275
Db 752327 GGACGCACTTGGCTTGGTGGAGCGCGCGT-----GTCAACTACGAACTCACC 752377
Qy 276 GlyAspLeuLysThrValGlyArgPheAspGlyGlnLeuGluSerThrMetIle 295
Db 752378 GATGAGCTGGACACCGTGGACCTGTGACTTCACGCGCACCTGCGCGGGTTACACC 752437
Qy 296 AlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValVal 315
Db 752438 GCCCATCCGACGTCGATCCGACACCGGTGAACCTGACGCGGTCTCTACTCTGTCGCC 752497
Qy 316 SerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSerProAspVal 335
Db 752498 CGCGACACAGAGTGCAGTACTCTGCTGATCGGCACCGACCGACCGTCTGTCGACGGT 752557
Qy 336 GluIleGlnLeuAspGlnProThrMetMetHisAspPheAlaIleThrGluAsnPheVal 355
```

## RESULT 2

US-09-103-840A-1  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.



```

; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 4,238-39 Length: 4411529
Score: 494.50 Matches: 180
Percent Similarity: 40.18% Conservative: 94
Best Local Similarity: 26.39% Mismatches: 251
Query Match: 15.70% Indels: 157
DB: 27

US-09-758-269-6 (1-599) x US-09-103-840A-1 (1-4411529)
QY 1 MetalaserPheThrAlaThr---AlaAlaValSerGly-----Arg 13
DB 749677 ATGGCCAGGAGTCGGGACCGGATCGGCGGTCTCGGGACCGTCAGCGGCATCGAAC 749736.
QY 14 TrpLeuGlyGlyAsnHisThrGlnProProLeuSerSerSerGlnSerSerAspLeuSer 33
DB 749737 TGGCGTAGCCTTCTCGGCTATGGCGGATCAGTCCGCGCATCCCGCGAAATGGGTG 749796
QY 34 TyrCysSer-----SerLeuProMetAlaSerArgValThrArgLeuAsnVal--- 50
DB 749797 TACACCGGCATCGTCAGGTGCTGTCGGCGGCCACCTTCGGGTCTGCAGCGCGTCG 749856
QY 51 -----SerSerAlaLeuHisThrProProAlaLeuHisPheProLysGlnSerSer 67
DB 749857 GCGCGGTGATCTCGAGCAGTCGCACGCGCGGTGC-----AGCAGCTCGTCG 749904
QY 68 AsnSerProAlaIleValVal-----LysProLysAlaLysGluSerAsn 82
DB 749905 CGAACACCGGTCTCGGAGTCTATCTGTCATGTTCTCACCAAGGCGGTACCGTTCCAAT 749964
QY 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaLagluGly 102
DB 749965 ATCAGTGAATAACCAATGTTATAGGAGATCGGCATGCCACCGCACCAAGCCCGAA--- 750021
QY 103 PheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerValGln 122
DB 750022 -----TCCAAACCCATATCTCGAG 750042
QY 123 IleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuProValVal 142
DB 750043 -----GGCTTCTGCGCGGTGAGCACCAGGTAACCTGCCACCGACCTGCGGTACC 750096
QY 143 GlyLysLeuProAspSerIleLysGlyValThrValArgAsnGlyAlaAsnProLeuHis 162
DB 750097 GCGCGCATTCGGAACACCTCGACGGCGGTATCTGCGTAAACGGCCCAACCGCGTCCG 750156
QY 163 Glu-----ProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaVal 180
DB 750157 GAGGTGACCGCGCCACC---TACCACCTGTTTCACGGCGGACGCCATGTCAGCGAGTC 750213
QY 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200
DB 750214 GCGGTGCGGACGGGAAGGCC-----CGCTGGTAT 750243
QY 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGlu----- 216
DB 750243 -----CTGTATAACAGGACTCGTGTACCGGGTCCAGCAGC 751275

```

```

DB 750244 CGCAATCGCTGGTCCGCACACCCCGCGGTGTCGGCGCGCTGGGGAGAGCCCATTTTCGGCC 750303
QY 217 ---LeuHisGlyHisThrGlyIleAlaAlaArgLeuMetLeuPheTyrAlaArgAlaAla 235
DB 750304 CGGCTCTACCCGCGCACCAGGATTATCAG----- 750333
QY 236 GlyIleValAspProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsn 255
DB 750334 -----GGCGGTCCCAACACCGATGCTGACCCACGCC 750366
QY 256 GlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsn 275
DB 750367 GGACGACCCCTGGCTTGTGAGCCGCGGTG-----GTCACTACGAATCACC 750417
QY 276 GlyAspLeuLysThrValGlyArgPheAspGlyGlnLeuGluSerThrMetIle 295
DB 750418 GATGAGCTGGACACCGTGGACCCCTGTCATTCGACGGCACCCTGCACGGCGTTACACC 750477
QY 296 AlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValVal 315
DB 750478 GCGCATCGCAGCGGTATCCGCACACGCGTGAATCGACGCGGTCTCTACTCTGTTCCGCC 750537
QY 316 SerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSerProAspVal 335
DB 750538 CGCGGACACAGAGTCAGTACTCGGTGATCGGCACCGACGACACGCTGTCGACGCTT 750597
QY 336 GluIleGlnLeuAspGlnProThrMetHisAspPheAlaIleThrGluAsnPheVal 355
DB 750598 GATATCGAGGTGGCGGATCCCGATGATGACAGCTTCTCCCTGACCGACAACTACGTG 750657
QY 356 ValValProAspGlnGlnValValPhe----- 364
DB 750658 GTGATCTACGACTGCTCGGTGACCTTCAGCCAAATGACAGTGGTCCGCGCGTCCGTGCCA 750717
QY 365 -----LysLeu 366
DB 750718 CGTGGTGTCAACGCGCCCGCAGGTGTGTGATCCAGTCCGTCTGGGCGGTCTCGGCATC 750777
QY 367 ProGluMetIle-----ArgGlyGlySer-----ProVal 376
DB 750778 CCGACCCGATAGCGCGGTGGGCAACCGGATGAGGATCTCGATCGCTCCCGTAC 750837
QY 377 ValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSer 396
DB 750838 GCCTGGAAACCCAGCTACCCGCGCGCTCGGTGTGTCGCGCGCGAGGT---GGCAAC 750894
QY 397 SerAsnIleLysTrpIleAspAlaProAspCysPheHisLeuTrpAsnAlaTrp 416
DB 750895 GAGGACGTGCGGTGTTGCATCATCGAACCCCTGCTACGTATACCCACCTTAACCCCTAC 750954
QY 417 GluGluProGluThrAsp---GluValValIleGly-----SerCysMet 431
DB 750955 TCGAGTCCCGAAGCGCGCTGAGTGTGTGTGGAGTGGTGGCTACTACTACCGATG 751014
QY 432 ThrProProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGlu 451
DB 751015 TTTGATCGCAGCCGCGGGTCCCGCGGTGACAGC-----CGGCGCTCGGTGATCGC 751068
QY 452 IleArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAsp 471
DB 751069 TGGACCATCAACCTGGCAGCGGTGCGGTGACCGCGAA-----TGCGCGCAGCAT 751119
QY 472 GlnGlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLys 491
DB 751120 CGGGG-----CAGGAGTTTCCCGCATCAACGAGACTCTGGTGGGTGGCGCGCATCGC 751173
QY 492 PheAlaTyr-----LeuAlaLeuAla 498
DB 751174 TTCGCTACACCGTCGGCATCGAGGTGGTTCCTCGTCGCGCGCGCGGTGCGTGTGTCG 751233
QY 499 GluProTrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLys 518
DB 751234 ACTCCG-----CTGTATAACAGGACTCGTGTACCGGGTCCAGCAGC 751275

```

```
Qy 519 LysHisLeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGly 538
Db 751276 GTCCGCTCGCTCGATCCCGACCTGCTGATCGCGGAGATGGTGTTCGTGCCGAACCCGCTCG 751335
Qy 539 GlyGlu---GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTriPlys 557
Db 751336 CGCGGTGCAGAGATGACGGGATTCATGGCTACGGCTGGCACCAGCGCGCGGACGAA 751395
Qy 558 SerGluLeuGlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuPro 577
Db 751396 GCGCAGCTGCTCTGCTGGATGCCAGACTCTCGAGTCGATCGGCACCGTCGACTGCCA 751455
Qy 578 SerArgValProTyrGlyPheHisGlyThrPhe-----IleGly 590
Db 751456 CAGCGTGTGGCGATGGGCTTCACGGCAACTGGGCGCGCACCTGACGGCGCTCGGG 751515
Qy 591 AlaAsp 592
Db 751516 TCGCAT 751521

RESULT 3
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2

Alignment Scores:
Pred. No.: 5,52e-37 Length: 4403765
Score: 475.50 Matches: 165
Percent Similarity: 40.35% Conservative: 90
Best Local Similarity: 26.11% Mismatches: 255
Query Match: 15.10% Indels: 123
DB: 3 Gaps: 23

US-09-758-269-6 (1-599) x US-09-103-840A-2 (1-4403765)
Qy 17 GlyAsnHisThrGlnProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSer 36
Db 1018941 GCGCGCTACCCATCAACCCGAGGAGCTGATCGGTGGCGGACATCCGG-----TCG 1018888
Qy 37 SerLeuProMetAlaSerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThr 56
Db 1018887 GGGCTTCCGGTGTGGGATGTGCTCGACGCGGCCAACAGGTGAGCGGCATTGCCGGTG 1018828
Qy 57 ProProAlaLeuHisPheProLysGlnSerSerAsnSerProAlaIleValLysPro 76
Db 1018827 ACTATCAGTCGAGNACCCGAGCGGTTCGCGACGCTGAATTTCGGTGGCAGCAGGCCA 1018768
Qy 77 LysAlaLysGluSerAsnThrLysGlnMetAsnLeuPheGlnA-gAlaAlaAla 96
Db 1018767 CCACGGTCAGTTTCG-----TTGTCAGCACCCACAC 1018738

97 LeuAspAlaAlaGluGlyPheLeuVal-SerHisGluLysLeuHisProLeuProLys-- 115
1018737 GAGGATCGTGAATGGACATCATCAGATTGTCCGMAGTACTTGTGCACCCCTCCCGAAGAC 1018678
116 -ThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProVa 135
1018677 GACGACCACCCCTATCGCACCGGTCCGTGGCGCACACAGACCACCGAATGGATGCCGAC 1018618
135 lArgArgAsnLeuProValValGlyLysLeuProAspSerIleLysGlyValTyrValAr 155
1018617 GACCTG-ACC---ACCCTGACAGCGGAAGTCCCGCCGACCTGAGACGCGCATCTACCTCG 1018562
155 qAsnGlyAlaAsnProLeuHisGluProValThrGlyHisPheHisPheAspGlyAspGl 175
1018561 CAACACGAGAACCCGCTACACCGCGCATTCGCGACCTACCCCTTCGATGGCGACGG 1018502
175 yMetValHisAlaValLysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGl 195
1018501 CATGATCCATGTCTCGCTTCGCTGATGGAAGACCTTCTACCGCAACGATTATTTCG 1018442
195 nThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGl 215
1018441 CACCGATGGATTCTTGCCGAGAACGAGCGCGCGCGCTGTGGCGC----- 1018393
215 yGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArg----- 232
1018392 -----GGTCTGGCAGAACCGGTGCACCTGCGCAAGCGGCAAGCGGAAACACGG 1018352
233 ----AlaAlaAlaGlyIleValAspProAlaHisGlyThrGlyVal---AlaAsnAlaGl 250
1018351 CTGGGGCGCTCGTGGCTCATGAAGGACGCGTGAGCACCACGCGTATCGTCCACCGAGG 1018292
250 yLeuVal-----TyrPheAsnGlyArgLeuLeuAlaMetSerGluAspAspLe 266
1018291 TATCGCGCTGACTAGCTTCTACCGAGTGGCGGATCTGTATCGGATCGAC----- 1018243
266 uProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAspPh 286
1018242 -CGTACTCGGCCNATACGGTC-----GGCAAGGAGAGTGGCAGCAAGAGTTTCGTT 1018190
286 eAspGlyGlnLeuGluSerThrMetIleAlaHisProLysValAspProGluSerGlyGl 306
1018189 CGACTGGGGCGGTGCG-----GCACATCCGAAGGTAGACAAACAGACCGCGCA 1018142
306 uLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSe 326
1018141 ACTGTCTTCTTCAACTACAGC---AAGCAAGAGCCGTATATGCGCTACGCGGTGTGCGA 1018085
326 rProAspGlyThrLysSerProAspValGluIleGlnLeuAspGlnProThrMetMetHi 346
1018084 CCAGAACAAATGAGCTTGTGCACTATGTGATGTTCGCTGCCCGGGCGCGGTACCGCA 1018025
346 sAspPheAlaIleThrGluAsnPheValValValProAspGlnGlnValValPheLysLe 366
1018024 TGACATGGCGTTCACCGAAATACGTAATCTCAACGAT----- 1017985
366 uProGluMetIleArgGlyGlySerProValValTyrAspLysAsnLysVal----- 383
1017984 -----TTTCCACTGTTTCGGATCCCGAGGTGCTCGAGCGCGCA 1017947
384 -----AlaArgPheGlyIleLeuAspLysTy 392
1017946 GTGCACCTACCAGCTTCTATCCGGAGATTCATCTCGGTTTCGCTGTT----- 1017895
392 rAlaGluAspSerSerAsnIleLysTrpIleAspAlaProAspCysPheCysPheHisLe 412
1017894 -GCTCGCGGAGGCAACAGACATCTCGTGTTCGAAGCGCATCAACGTTTCGTTGTCACCT 1017836
412 uTrpAsnAlaTrpGluGluProGluThrAspGluValValIleGlySerCysMetTh 432
1017835 CACCAACGCTTACGAG-----CAGGCGCAGAGATCTGCTCGACGCG----- 1017793
432 rProAspSerIlePheAsnGluSerAsp----- 442
```

```
Db 101792 -----TTCTACGAGCGCATCGCAGCCACTTGCACCGAGGACGAA 1017749
Qy 443 -----GluAsnLeuLysSerValLeuSerGlu1 452
Db 101748 GTGGGAGAGCTGTTTCGGTTCCTGCTCGATCGCTGCTACCGCTACATCGGTG 1017689
Qy 452 eArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspG1 472
Db 1017688 GCGGCTCAATATGGTGACCGG-----GCAGTCCACGAGGA 1017653
Qy 472 nGlnValAsn-----LeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgly 489
Db 1017652 GCACCTGTCGAGTCCATCACCAGGTTCCGAACCATCAACCGGATTCACCGCCAGCAG 1017593
Qy 489 sThrPheAlaTyrLeuAlaLeuAlaGluPro---TrpProLysValSerGlyPheAl 508
Db 1017592 CTACCGTTACACTATGCGCTACCGGCAACCGAGTTGTTCTGTTCCGACGGACTGGT 1017533
Qy 508 aLysValAspLeuThrThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrG1 528
Db 1017532 TAAGCACGATCTGCTCACCGGAACACACGAGTGTTACTGTTCTGCGTGACGGCTCTACGG 1017473
Qy 528 yGlyGluProLeuPheLeuProGlyGluGlyGlu---GluAspGluGlyTyrIleLe 547
Db 1017472 AGTGAGACCGGATGGCTCCACGGGTGGGACGACGCGCGGAGGACGCGCTATCTGGT 1017413
Qy 547 uCysPheValHisAspGluLysThrTriPlysSerGluLeuGlnIleValAsnAlaValSe 567
Db 1017412 CACCCTCACCGACATGAACGACGACGATCGTATTGCTGGTTTTCGACCGCGCGG 1017353
Qy 567 rLeuGluValGluAlaThrValLys-----LeuProSerArgValProTyrGlyPheHi 585
Db 1017352 CCCCAGCGATGCGCCGATATGCAAGCTTGCACTGCCGGAACGTATTTCCAGCGGCACGCA 1017293
Qy 585 sGlyThrPheIleGlyAlaAspLeuAlaLys 596
Db 1017292 TTCGGCGTGGTGGCGGCGCGCGGATGCGTGC 1017259

RESULT 4
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 441529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 5,53e-37 Length: 4411529
Score: 475.50 Matches: 165
Percent Similarity: 40.35% Conservative: 90
Best Local Similarity: 26.11% Mismatches: 255
Query Match: 15.10% Indels: 123
DB: 3 Gaps: 23

US-09-758-269-6 (1-599) x US-09-103-840A-1 (1-4411529)

Qy 17 GlyAsnHisThrGlnProLeuSerSerGlnSerSerAspLeuSerTyrCysSer 36
```



16390	-----GAAC	TGGATCCTCGT	---ACACTTGA	AAACTCGCGGACAC	TTTGCAC	TACACGCGC	16344
289	GlnLeu	GluSer	---ThrMet	IleAlaHisPro	LysValAsp	ProGluSerGly	GluLeu 307
16339	CAAGTTAC	CGCAAAACCCAC	CGCCCATCC	CAAAATATG	ACCCGAAAC	CGGTGACTTG	16280
308	PheAla	LeuSerTyr	AspValValSer	LysProTyr	LeuLysTyr	PheArgPheSerPro	327
16279	TTG	---	---	---	---	---	16259
328	AspGly	ThrLysSerPro	AspVal	---	---	---	335
16258	AAGGGC	GAACCACTCC	GACATGGCCT	TATTAC	TTCGAC	ACGCGCAAGGTGACA	16199
336	---Glu	IleGlnLeu	AspGlnPro	-----ThrMet	MetHisAsp	PheAlaIleThrGlu	352
16198	CATGAA	ACTTTGGTTTG	CAGCAGCCCT	TATGCG	CATTATG	CGCATTTGCCAT	TACCGA 16139
353	Asn	PheValValPro	AspGlnGlnVal	ValPheLysLeu	ProGluMet	IleArgGly	372
16138	AAT	TGTCATTTTCC	CAATATATG	CGCGGC	CAACAAG	CCTG	16082
373	GlySer	ProVal	---ValTyr	AspLysAsnLysVal	AlaIleArgPhe	GlyIleLeuAspLys	391
16081	AAACAG	CAATTTATAT	TATGTTGG	GCAGCCGA	ACTTGGG	CAGCTCATTTGGG	TACTC----- 16028
392	TyrAla	GluAspSerSer	AsnIleLysTrp	IleAspAlaPro	AspCysPheCysPhe	His 411	
16027	CGCGCG	CGCAGGCAGT	CTGATTCTG	CTGGCTCA	AGGCACCGCG	CTCTGGTATTT	TCAT 15968
412	LeuTrp	AsnAlaTrp	GluGluPro	GluThrAsp	GluValValIle	GlySerCysMet	431
15967	GTTGT	GAATCTTTGGG	AACTCGGA	ACCAAGATTTAT	TATCGAC	CTTTATG	GAAGTGAATC 15908
432	ThrPro	-----Pro	AspSer	-----	-----	-----	436
15907	CTCGGT	CTCCCTTCC	CAACTCACA	AAACCA	CCCTTCG	CCCTCGAGAA	CCGTACCA 15848
437	-----	-----	-----	-----	-----	-----	15847
15847	CGCCT	GACTCGTTGG	GAATATG	ACTCTGAT	AGCAGCG	CAGCATCA	AGGGAACCCGG 15788
449	LeuSer	GluIleArg	LeuAsnLys	ThrGlyGluSer	Thr	-----	462
15787	CTACAG	ATTTCTTTG	CGAAATGCC	AATCATG	GATTTCTTC	TGTCG	CCCTGCAATGCAAC 15728
463	-----	-----	-----	-----	-----	-----	462
15727	CGTAT	GGCTTTATG	GGGTGG	CGATCC	ACCAACCA	CTTTGG	CCATCATCAGCAGCCGAG 15668
473	GlnVal	AsnLeuGlu	AlaGlyMet	ValAsnArg	AsnMetLeu	GlyArgLysThr	LysPhe 492
15667	AAATA	---	---	---	---	---	15659
493	AlaTyr	LeuAlaLeu	AlaGluPro	TrpTrpLysVal	SerGlyPhe	AlaLysValAspLeu	512
15658	GCGTACA	ACTCACTCGC	---ATCTGG	-----	-----	-----	15629
513	ThrThr	GlyGluValLys	LysHisLeuTyr	GlyAspAsnArg	TyrGlyGlyGlu	ProLeu	532
15628	CACCGAG	GTGACTAC	GACCTCTGG	TACTCGG	CGAAGCTCGG	CGCCAGCGG	CCGAGCCGCGCC 15569
533	PheLeu	Pro	---Gly	GluGlyGlyGlu	AspGluGlyTyr	IleLeuCysPheValHis	551
15568	TTGTC	CTCTAG	AAGTCCG	ACCGCCCG	CAAGGTGAC	CGGTACTT	GCTGACCGTGGTGGT 15509
552	AspGlu	LysThrTrp	LysSerGluLeu	GlnIleValAsn	AlaValSerLeu	Glu	----- 569
15508	CGCCT	CGATGAAA	ATCGCAG	CGATCTGG	TAAATCTCG	ACACTCA	AGACATCCAGCTCGGT 15449
570	ValGlu	AlaThrVal	LysLeuPro	SerArgValPro	TyrGlyPhe	HisGlyThrPhe	Ile 589
15448	CCGCTG	GCACCATCA	AGCTGCTG	CACTTCGG	GTCTCAAG	CGCGCTCTCA	TGGCTGGTGGTA 15389

**RESULT 6**

```
US-08-976-063E-21/c
; Sequence 21, Application US/08976063E
; Patent No. 6524831
; GENERAL INFORMATION:
; APPLICANT: Steinhuchel, Alexander
; APPLICANT: Priefert, Horst
; APPLICANT: Rabenhorst, Jurgens
; TITLE OF INVENTION: SYNTHETIC ENZYMES FOR THE PRODUCTION OF CONFERYL
; TITLE OF INVENTION: ALCOHOL, CONFERYLALEHYDE, FERULIC ACID, VANILLIN AND
; FILE REFERENCE: Baver-9998-CAO
; CURRENT APPLICATION NUMBER: US/08/976,063E
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 196 49 655.1 GERMANY
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 1518
; TYPE: DNA
; ORGANISM: not required under old rule
US-08-976-063E-21
```

Alignment Scores:			
Pred. No.:	9,53e-36	Length:	1518
Score:	411.50	Matches:	137
Percent Similarity:	40.68%	Conservative:	79
Best Local Similarity:	25.80%	Mismatches:	192
Query Match:	13.06%	Indels:	123
DB:	4	Gaps:	19

  

```
US-09-758-269-6 (1-599) x US-08-976-063E-21 (1-1518)
```

Qy	118 AspProSerValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArg	137
Dd	1497 GACCGC-----CAATTAGTAGGAACACTTCTCCACCACGGTATAGAGGCAGACTTGTTTC	1444
Qy	138 AsnLeuProValValGlyLysLeuProAspSerILELYSGlyValTYrValArgAsnGLY	157
Dd	1443 GATCTAGAGGTTCAGCGCGAAATCCCCAAAATCAATAAATGGAACGTTCTACCGCTAATACG	1384
Qy	158 AlaAsnProLeuHisGluProValThrGlyHisPhePheAspGlyAspGlyMetVal	177
Dd	1383 CcAGAGCcTCAAGTTAcCCcACAaaaATTCcAcCCTTCATgATGCAGATGGAATGGCC	1324
Qy	178 HisAlavalysPheGluHisGlysSerAlasetyrAlacyeArgPheThrglinThrAsn	197
Dd	1323 TcTgcCTccAcCTTCGaAgATGtCATGTgcACTTCATcAGTCGCTGGGTAAAAACCGCT	1264
Qy	198 ArgPheValGlnGluArgGlnLeuGlyArgProValPheProLysAlalileGlyGluLeu	217
Dd	1263 CGAtTCAGCGCGCAACGACTAGCGCAAATTCGCTATTTGGCATGTATCACAAACCCCTAT	1204
Qy	218 HisGlyHisThrGlyTleAlaalargLeuMetLeuPheTyralaaAlaAlaGlylle	237
Dd	1203 ACCGACGACcACcAGTGTAAAGGAcTA-----	1177
Qy	238 ValAspProAlahisglyThrGlyValalaAsnalaglyLeuValtyrPheAsnGlyArg	257
Dd	1176 -----GACCGCACCGTGTGCCAATACAAGCATCATATGACCATCACGCAAG	1132
Qy	258 LeuLeuAlaMetSerCluhaspLeuProtyrGlnValGlnIleThrProAsnGlyAsp	277
Dd	1131 GTGTCGGCGGTGAAGAAGACGCCCTACCGTAC-----GAATGGATCCTCGT---ACA	1081
Qy	278 LeuLystrValiclyArgPheAspPheAspGlyGinLeuGluSer---ThrMetIleAla	296
Dd	1080 CTTGAAACTCGGGACACTTCGACTACGACGGCCAAGTTACCAGGCCAAACCCACACCGCC	1021
Qy	297 HisProLyseValaspProGluseryGlyGluLeuPheAlaLeuSertyrAspValValser	316



Db 215 CACAGTTTGTACTTTAAAGAGGACATGTACATACCACAGAGGTTTCATCCGCACTGAT 274  
Qy 198 ArgpheValGlnGlu-----ArgGlnLeuGlyArgProVal 209  
Db 275 GCTTACGTACGGGCAATGACTGAGAAAGGATCGTCATAACAGAAATTTGGCACCCTGTGCT 334  
Qy 210 PheProLysAlaIleGlyGluLeuHisGlyHisThrGlyLeuAlaArgLeuMetLeuPhe 229  
Db 335 TTTCCAGATCCCTGCAAGAAATATA-----TTTCCAGGTTTTTTTCTTAC 379  
Qy 230 TyrAlaArgAlaAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAlaAla 249  
Db 380 TTC-----CGAGGAGTGGAGGTACTACCAAT 406  
Qy 250 GlyLeuValTyPhe-----AsnGlyArgLeuLeuAlaMetSerGluAspLeu 266  
Db 407 GCCCTTGTGTTAA-TATCTACCCAGTGGGGAAGATTACTATGCTGCACAGACCAACTT 465  
Qy 267 ProTyrglnValGlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAspPh 286  
Db 466 CATTACAAAGGTT-----AATCCTGAG--ACCTTGGAAACAATTAAGCAGGTTGACCT 516  
Qy 286 eAspGlyGlnLeuGlu---SerThrMetIleAlaHisProLysValAspProGluSer-- 304  
Db 517 TTGCAACTATGTCTCAGTCAATGGAGCCACTGCTCACCCACATTTGAAATGATGGGAC 576  
Qy 305 -----GlyGluLeuPheAlaLeuSerTyAspValIse 316  
Db 577 TGTTTACCAATTTGGTAATGCTTTGGGAAAAATTTTCAATTCCTACAAATTTGTA 636  
Qy 316 rlyAsProTyrglnLeuLysTyPheArgPheSerProAspGlyThrLysSerProAspValG 336  
Db 637 GATCCCACTACCTACCAAGCAGACAGCAAGAGATCCA-----ATAAGCAAGTCAGAGATCGT 690  
Qy 336 uileGlnLeuAsp-----GlnProThrMetMetHisAspPheAlaIleTh 351  
Db 691 TGTACAAATCCCTCGAGTCGAGTCAAGCCATCTTACGCCATATGTTTGGTTGATG 750  
Qy 351 rGluAsnPheValValProAspGlnGlnValPheLysLeuProGluMetIleAr 371  
Db 751 TCCCAACTATATGTTTTTGTGGAGACACCACTCAAAATTAATCTGTTCAAGTTCTTTC 810  
Qy 371 gGlyGlySerProValValTyAspLysAsnLysValAlaArgPhe----- 386  
Db 811 TTCTGAGGT-----CTTGGGAGCCAAATACATGATGTTTGTGAATCCAATGAAC 864  
Qy 387 -----GlyIleLeuAspLysTyrglnAlaGluAspSerSerAsnIleLysTr 401  
Db 865 CATGGGGTGTGCTTTCATATTTGTCACAAAAAAGAAAGAAAGTATATCAATAATAATA 924  
Qy 401 pIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluProGluTh 421  
Db 925 CAGGACCTCTCT--TTTAACTCTTTCATCATCAATCACTATGAAGACCATGAT 981  
Qy 421 rAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsn----- 439  
Db 982 T-----CTGATGTGGATCTCTGTTGCTGGAAGGATTTGAATTTGTTTATAATTTT 1035  
Qy 440 -----GluSerAspGluAsnLeuLysSerValLeuSerGluIleAr 453  
Db 1036 ATATTAGCAATTTACGTGAGAACTGGGAAGAGGTGAAAAA--ATGCCACAGAGGCT 1093  
Qy 453 gluAsnLeuLysThrGlyGluSer-----ThrArgAr 464  
Db 1094 CTTACGCTGAATTTAGGATACGTACTTCTTTGAATATTGACAAAGCTGACACAGGC 1153  
Qy 464 gProIleLeuSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMetValAsnArgAs 484  
Db 1154 AAGAATTTAGTCACACTCCCAACACAACTGCCCTCAATGCAATTTCTGTGAGTGACGAGACC 1213  
Qy 484 nMetLeuGlyArgLysThrLysPheAlaTyrlLeuAlaLeuAlaGluProTrpProLysVa 504  
Db 1214 ATCT---GGCTGGAACCTGAGGTCTCTTTTCAGGGCCTCGCCCAAGCATTTGAGTTTCTCT 1270

Qy 504 IserGlyPheAlaLys-----ValAspLeuThr-- 513  
Db 1271 CAAATCAATTTACCAGAGATATGGTGGGAAACCTTACACATATGCATATGACCTTGGCTTG 1330  
Qy 514 -----ThrGlyGluValLysLysHisLeuTy 522  
Db 1331 AATCACTTTGTCACAGACAGGCTCTGTAAAGCTGACGTCACAAACCTAAAGAAACCTGGGTA 1390  
Qy 522 rGlyAsp-----AsnArgTyrglyGlyGluProLeuPheLeuPro---GlyGluGlyG 539  
Db 1391 TGGCAA-CAGCCTGATTCATACCCCTCAGAACCTATCTTTTCTCACCAGATGCCTT 1449  
Qy 539 yGluGluAspGluGlyTyrlleLeuCysPheValHisAspGluLysThrTrpLys----- 557  
Db 1450 GGAGGAAGATGACGGTGTAGTCTGAGTGTGGTGAGCCCTGGGGCAGGACAAAGCC 1509  
Qy 558 -SerGluLeuGlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuPr 577  
Db 1510 TGCTTATCTTCTGATTTCTGAATGCCAAGGACTTGAGTGAAGTTGCCAGGCTGAAAGTGA 1569  
Qy 577 oSerArgValProTyrglyPheHisGlyThrPhe 588  
Db 1570 GATTAACATCCCTCCTCACCTTTTATGGACTGTT 1603

## RESULT 8

US-08-488-305A-1  
; Sequence 1, Application US/08488305A  
; Patent No. 5679772  
; GENERAL INFORMATION:  
; APPLICANT: B vik, Claes Olof; Eriksson, Ulf; Peterson, Per A.  
; TITLE OF INVENTION: Isolated Protein Receptors, Antibodies Which  
; TITLE OF INVENTION: bind Thereto, Nucleic Acid Sequence Coding  
; Patent No. 5679772  
; TITLE OF INVENTION: Therefor, And Uses Thereof  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/488,305A  
; FILING DATE: 7-JUNE-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kohli, Vineet  
; REGISTRATION NUMBER: 37,003  
; REFERENCE/DOCKET NUMBER: LUD 5280.3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2629 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHEITICAL: no  
; ANTI-SENSE: no  
US-08-488-305A-1

Alignment Scores: 5.23e-13 Length: 2629  
Pred. No.: 211.00 Matches: 130  
Score:

Percent Similarity: 40.03%      Conservative: 99  
Best Local Similarity: 22.73%      Mismatches: 215  
Query Match: 6.70%      Indels: 131  
DB: 1      Gaps: 26

US-09-758-269-6 (1-599) x US-08-488-305A-1 (1-2629)

Qy 105 ValSerHisGluLysLeuHisProLysThrAlaAspProSerValGlnIleAla 124  
Db : : : : :  
Qy 14 ATGTCCAGCCAAAGTTGAACATCCA-----GCT 40  
Db : : : : :  
Qy 125 GlyAsnPheAlaProValAsnGlnProValArgAsnLeuPro----- 140  
Db : : : : :  
Qy 41 GGTGGTTACAGAAACCTGTTGAACCTGTGAGGAACTATCTCACCGCTCACAGCCCAT 100  
Db : : : : :  
Qy 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160  
Db : : : : :  
Qy 101 GTTACAGCAGGATCCCTCTGGCTAACCCGCGAGTCTCTCGATGTTGGGCCAGGACTC 160  
Qy 161 LeuHis-----GluProValThrClyHisHisPheAspGlyAspGlyMetVal 177  
Db : : : : :  
Qy 161 TTTGAGTTGATCGGAACCAATTT-----TACCACCTGTTTGTGGCAAGCCCTCCTA 214  
Db : : : : :  
Qy 178 HisAlaValLysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsn 197  
Db : : : : :  
Qy 215 CACAAGTTTGACTTTAAGAAGGACATGTCATACACACAGAGTTTCATCCGCACTGAT 274  
Qy 198 ArgPheValGlnGlu-----ArgGlnLeuGlyArgProVal 209  
Db : : : : :  
Qy 275 GCTTACGTACGGCAATGACTGAGAAAAGGATCGTCATAACAGAAATTTGGCAGCTGTCT 334  
Qy 210 PheProLysAlaIleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPhe 229  
Db : : : : :  
Qy 335 TTCCAGATCCCTGCAAGAAATA-----TTTCCAGGTTTTTTTCTTAC 379  
Qy 230 TyrAlaArgAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsnAla 249  
Db : : : : :  
Qy 380 TTC-----CGAGGAGTGGAGTTTACTGACAA 406  
Qy 250 GlyLeuValTyrPhe-----AsnGlyArgLeuLeuAlaMetSerGluAspLeu 266  
Db : : : : :  
Qy 407 GCCCTTGTTAA-TATCTACCCAGTGGGGAAGATTACTATGCTGCACAGAGACCACTT 465  
Qy 267 ProTyr-GlnValGlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAspPh 286  
Db : : : : :  
Qy 466 CATTACAAAGGTT-----AATCCTGAG---ACCTTGGAAACAATTAAGCAGGTGACCT 516  
Qy 286 eAspGlyGlnLeuGlu---SerThrMetIleAlaHisProLysValAspProGluSer-- 304  
Db : : : : :  
Qy 517 TTGCAACTATGCTCAGTCAATGGAGCCAGTCTCACCCCAATTTGAAATGATGGAC 576  
Qy 305 -----GlyGluLeuPheAlaLeuSerTyrAspValValSe 316  
Db : : : : :  
Qy 577 TGTTTACAACATTTGTAATTGCTTTGGGAAAAATTTTCAATTGCTTACAATATTGATAA 636  
Qy 316 rLysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSerProAspValG 336  
Db : : : : :  
Qy 637 GATCCCACTTACAAAGCAGACAGCAAGAGATCCA-----ATAAGCAAGTCAGAGATCGT 690  
Qy 336 uileGlnLeuAsp-----GlnProThrMetMetHisAspPheAlaIleTh 351  
Db : : : : :  
Qy 691 TGTACAATTTCCCTCAGTGCAGGATTCAGCCATCTTACGTCATAGTTTGTGTTGAC 750  
Qy 351 rGluAsnPheValValValProAspGlnGlnValValPheLysLeuProGluMetIleAr 371  
Db : : : : :  
Qy 751 TCCCAACTATATTGTTTGTGGAGACACACAGTCAAAATTAATCTGTTCAAGTTTCTTTC 810  
Qy 371 gGlySerProValValTyrAspLysAsnLysValAlaArgPhe----- 386  
Db : : : : :  
Qy 811 TTATGAGAGT-----CTTTGGGAGCAATTTACATGGAATGTTTGAATCCAATGAAC 864  
Qy 387 -----GlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLysTr 401  
Db : : : : :  
Db 865 CATGGGGGTTGGCTTCATATTGCTGACAAAAAGAAAAAGTATATCAATAATAATA 924  
Qy 401 pIleAspAlaProAspCysPheCysPheHisLeuTyrAsnAlaTrpGluGluProGluTh 421  
Db : : : : :  
Qy 925 CAGGACCTCTCTCT---TTTAACTCTTTTCATCATCAATACCTATATGAACACCATGATT 981  
Qy 421 rAspGluValValIleGlySerCysMetThrProAspSerIlePheAsn----- 439  
Db : : : : :  
Qy 982 T-----CTGATTGTGGATCTCTGTGCTGAAAGGATTTGAATTTGTTTATATATT 1035  
Qy 440 -----GluSerAspGluAsnLeuLysSerValLeuSerGluIleAr 453  
Db : : : : :  
Qy 1036 ATATTGACCAATTACCTGAGAACTGGGAAGAGGTGAAAAA---ATGCCAGAAAGGCT 1093  
Qy 453 gLeuAsnLeuLysThrGlyGluSer-----ThrArgAr 464  
Db : : : : :  
Qy 1094 CCTCAGCCTCAAGTTAGGATAGATACGTACTTCTTTGAATATTGACAGGCTGACACAGC 1153  
Qy 464 gProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMetValAsnArgAs 484  
Db : : : : :  
Qy 1154 AAGATTATGATCAGCTCCCAACACACTGCCACTGCAATTTCTGTCAGTGCAGGAC 1213  
Qy 484 nMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluProTrpProLysVa 504  
Db : : : : :  
Qy 1214 ATCT---GGCTGGAACCTGAGTTCTCTTTTTCAGGGCTCGCCAAAGCATTTGAGTTTCT 1270  
Qy 504 lSerGlyPheAlaLys-----ValAspLeuThr-- 513  
Db : : : : :  
Qy 1271 CAAATCAATTACCAGAAAGTATGTTGGGAAACCTTTACATATGATGACTTGGCTTG 1330  
Qy 514 -----ThrGlyGluValLysLysHisLeuTy 522  
Db : : : : :  
Qy 1331 AATCACTTTGTCAGACAGGCTCTGTAACTGAACGTCAAACTAAAGAAACCTGGTA 1390  
Qy 522 rGlyAsp-----AsnArgTyrGlyGluProLeuPheLeuPro---GlyGluGlyG 539  
Db : : : : :  
Qy 1391 TGGCAA-GAGCCTGATTACATACCCCTCAGAACTATCTTTGTTCTCACCAGATGCTT 1449  
Qy 539 yGluGluAspGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLys----- 557  
Db : : : : :  
Qy 1450 GGAGGAAGATGACGGTGTAGTTCTGAGTGTGGTGTGAGCCCTGGGGCAGACAAAGGCC 1509  
Qy 558 -SerGluLeuGlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuPr 577  
Db : : : : :  
Qy 1510 TGCTATCTTCTGATTCGATGCAAGGACTTGAGTGAAGTTGCCAGGGCTGAAGTGA 1569  
Qy 577 oSerArgValProTyrGlyPheHisGlyThrPhe 588  
Db : : : : :  
Qy 1570 GATTAACTCCCTGCTCACCTTTTCATGACTGTTT 1603

RESULT 9  
US-09-385-259-1  
; Sequence 1, Application US/09385259  
; Patent No. 6201114  
; GENERAL INFORMATION:  
; APPLICANT: Aguirre, Gustavo D.  
; APPLICANT: Acland, Gregory M.  
; APPLICANT: Ray, Kunal  
; TITLE OF INVENTION: IDENTIFICATION OF CONGENITAL STATIONARY NIGHT BLINDNESS  
; TITLE OF INVENTION: IN DOGS  
; FILE REFERENCE: 19603/2481  
; CURRENT APPLICATION NUMBER: US/09/385,259  
; CURRENT FILING DATE: 1999-08-30  
; EARLIER APPLICATION NUMBER: 60/103,219  
; EARLIER FILING DATE: 1998-10-06  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1724  
; TYPE: DNA  
; ORGANISM: Canis familiaris  
US-09-385-259-1



## Alignment Scores:

Pred. No.: 3,3e-13 Length: 1724  
Score: 210.00 Matches: 113  
Percent Similarity: 37.08% Conservative: 98  
Best Local Similarity: 19.86% Mismatches: 218  
Query Match: 6.67% Indels: 140  
DB: 3 Gaps: 25

US-09-758-269-6 (1-599) x US-09-385-259-1 (1-1724)

QY	120	SerValGlnIle	-----AlaGlyAsnProValAsnGluInProValArg	136
DB	31	TCCATCAAGTGGAGCATCCCGCGGGGTACAGAAGCTGTTGAACCGTGGAGAG	90	
QY	137	ArgAsnLeuPro	-----ValValGlyLysLeuProAspSerIleLysGlyVal	152
DB	91	CTGCTCGCGCTCACCGCCACGTCAGACGAGGAGTCCGCTCTGGCTCAGCGGCGT	150	
QY	153	TyrValArgAsnGlyAlaAsnProLeuHis	-----GluProValThrGlyHisHis	169
DB	151	CTCCTCCGATGGCGACCGGGCTCTTCGAGGTGGATCTGAACCATTT	204	
QY	170	PhePheAspGlyAspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyr	189	
DB	205	CTGTTTGCGGACACGCCCTCTGACACAGTTTCGACTTTAAGAAGGACACGTCACCTAT	264	
QY	190	AlaCysArgPheThrGlnThrAsnArgPheValGlnGlu	-----	202
DB	265	CACAGAAGGTTTCATCCGACCGATGCTTACGTCGGGCAATGACCGAGAAAGGATCGTC	324	
QY	203	---ArgGlnLeuGlyArgProValPheProLysAlaIleGlyLeuHisGlyHisThr	221	
DB	325	ATAACGGAATTTGGCACCTGTGCTTCCAGATCCCTCGAAGAATATA	372	
QY	222	GlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspProAla	241	
DB	373	-----TTTTCCAGGTTTTTTTCTTACTTC	396	
QY	242	HisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAlaMet	261	
DB	397	CGAGGAGTGGAGTCACTGACAAATGCCCTTGT	444	
QY	262	SerGluAsp	-----AspLeuProTyrGlnValGlnIleThrProAsnGly	276
DB	445	GGGGAAGATTACTACGCTCGACGGAGACCACTTCATTACAAAGATTAACTCCTGAG	501	
QY	277	AspLeuLysThrValGlyArgPheAspGlyGlnLeuGlu	---SerThrMetIle	295
DB	502	ACCCTGGAGACAAATTAGCAGGTTGATCTCTGCAACTACGTCCTGTCATCGAGCCACC	561	
QY	296	AlaHisProLysValAspProGluSer	-----GlyGlu	306
DB	562	GCTCACCCCCACATTGAAATATGATGGAGCTGTTTACACATTTGGTAATTTGGGAA	621	
QY	307	LeuPheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSer	326	
DB	622	AAATTTTTCGATTCGCTACAAATATTGTAAGAATCCCTCCACTCCAAGCAGACAAAGAAAT	681	
QY	327	ProAspGlyThrLysSerProAspValGluIleGlnLeuAsp	-----Gln	341
DB	682	CCA-----ATAAGCAAGTCCGAGGTCGTCGTACAAATCCCTCGACGACCGATTCAG	735	
QY	342	ProThrMetHisAspPheAlaIleThrGluAsnPheValValProAspGlnGln	361	
DB	736	CCATCGTAGTCCATAGTTTGGTTGACTCCCACTATATGTTTGTGGAGAGGCCA	795	
QY	362	ValValPheLysLeuProGluMetIleArgGlyGlySer	-----	374
DB	796	GTCAAAATTAACCTGCTCAAGTTCTCTTCTGTCGAGTCCTTGGGAGCCCACTACATG	855	
QY	375	-----ProValValTyrAspLysAsnLysValAla	384	
DB	856	GATTTGTTAGTCCAAATGAACCATGGGGTTTGGCTTCACATCGCTGACAAAAAAGA	915	

QY	385	ArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLysTrpIleAspAla	404
DB	916	AAAAAGTATCTCAATAAAGTACAGGACCTCTCTCTTAAATCTC	960
QY	405	ProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluProGluThrAspGluVal	424
DB	961	-----TTCCATCATATCAATCTTACGAGACAAATGAGTTT	999
QY	425	ValValIleGlySerCysMetThrProAspSerIlePheAsnGluSerAsp	443
DB	1000	ATTGTTGATCTCTGCTGCGGAAAGATTGAAATTCGCTACAAATCTTGTATTAGCC	1059
QY	444	AsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGluSerThrArg	463
DB	1060	AAATTTACGTGAGAACTGGGAAGAGGTGAAA	1110
QY	464	ArgProIleIleSerAsnGluAspGlnValAsnLeuGluAlaGlyMetValAsnArg	483
DB	1111	CAGCCTGAAGTTAGGAGATCCGCTCTCTTGAATATATCGACAGGCCGACACAGCAAG	1170
QY	484	AsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeu	497
DB	1171	AACTAGTACCTTCCCAACACGACGGCCACTGCAACTCTCGCGACGACGACCATC	1230
QY	498	---AlaGluProTyrProLysValSerGly	510
DB	1231	TGGCTGGAACCTGAGGTTCTCTCTCAGGGCTCTGCAAGCCTTTTGAGTTTCTCTCAAATC	1290
QY	511	AspLeuThrThr	523
DB	1291	AACTATCAGAAGTATGGCGGAAGCCTTACAGCTACGCTATGGACTTGGCTTGAATCAC	1350
QY	523	-----	523
DB	1351	TTGCTTCGGACAGCTCTGCAAGCTGAACGTCAAGACTAAAGAAACGTGGGTATGGCA	1410
QY	524	---AspAsnArgTyrGlyGlyProLeuPheLeuPro	541
DB	1411	GAGCCGACTCATACCCATCAGAACCCATCTTTGTTCTCACCCAGATGCTTTGGAAGA	1470
QY	542	AspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLys	559
DB	1471	GATGATGTGTAGTTCTGAGTGTGTGTGAGCCCTGGGCGAGGACAAAAGCCTCTTAT	1530
QY	560	LeuGlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArg	579
DB	1531	CTTCTGATTTGAAATGCCAAGATTGAGTGAAGTTGCCAGGCTGAAGTGGAGATTAAAC	1590
QY	580	ValProTyrGlyPheHisGlyThrPhe	588
DB	1591	ATCCCTGTCACTTTTCATGGAGCTGTC	1617

## RESULT 10

US-09-645-370-1  
; Sequence 1, Application US/09645370  
; Patent No. 6428958  
; GENERAL INFORMATION:  
; APPLICANT: Aguirre, Gustavo D.  
; APPLICANT: Acland, Gregory M.  
; APPLICANT: Ray, Kunal  
; TITLE OF INVENTION: IDENTIFICATION OF CONGENITAL STATIONARY NIGHT BLINDNESS  
; FILE REFERENCE: 19603/2481  
; CURRENT APPLICATION NUMBER: US/09/645,370  
; PRIORITY FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 09/385,259  
; PRIOR FILING DATE: 1999-08-30  
; PRIOR APPLICATION NUMBER: 60/103,219  
; PRIOR FILING DATE: 1998-10-06  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 1

```
; LENGTH: 1724
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-645-370-1

Alignment Scores:
  3.3e-13      Length: 1724
Pred. No.:    210.00    Matches: 113
Score:         37.08%   Conservative: 98
Best Similarity: 19.86% Mismatches: 218
Query Match:    6.67%   Indels: 140
DB:             4       Gaps: 25

US-09-758-269-6 (1-599) x US-09-645-370-1 (1-1724)
Qy 120 SerValGlnIle-----AlaGlyAsnPheLaProValAsnGluInProValArg 136
Db 31 TCCATCCAAAGTGGAGCATCCCGCGCGGTTTACAAAGAGCTGTTTGAACCGTGAAGAG 90
Qy 137 ArgAsnLeuPro-----ValValGlyLysLeuProAspSerIleLysGlyVal 152
Db 91 CTGTCGTCGCGCTCACCGCCACGCTGACAGCGAGGATCCCGCTCTGCGCTCACGGGCGAGT 150
Qy 153 TyrValArgAsnGlyAlaAsnProLeuHis-----GluProValThrGlyHisHis 169
Db 151 CTCCTCCGATCGGACCGCGGCTCTTCGAGGTTGGATCTGAACCATTT-----TACCAC 204
Qy 170 PhePheAspGlyAspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyr 189
Db 205 CTGTTTGACGACAAAGCCCTTCGACAAAGTTCGACTTTAAAGAAGGACACGCTCACCTAT 264
Qy 190 AlaCysArgPheThrGlnThrAsnArgPheValGlnGlu----- 202
Db 265 CACGAAGAGTTTCATCCCGCACCGATGCTTACGTCCGGGCAATGACCGAGAAAGATCGTC 324
Qy 203 ---ArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHisThr 221
Db 325 ATAACGGAATTTGGCACCTGTGCGTTCACAGATCCCTCGACAGATATA----- 372
Qy 222 GlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaGlyIleValAspProAla 241
Db 373 ---TTTTCCAGGTTTTTTTCTTACTTC----- 396
Qy 242 HisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAlaMet 261
Db 397 CGAGGAGTGGAGGTCACTGACAATGCCCTTGT-----AACGTCTACCCAGTA 444
Qy 262 SerGluAsp-----AspLeuProTyrGlnValGlnIleThrProAsnGly 276
Db 445 GGGGAAGATTACTACGCTCGCAGGAGACCACTTCATTACAAAGATTAACTCTGAG--- 501
Qy 277 AspLeuLysThrValGlyArgPheAspPheAspGlyGlnLeuGlu---SerThrMetIle 295
Db 502 ACCCTGGAGACAAATTAAAGCAGTTGATCTCTGCAACTACGCTCTGTCAATGGAGCCACC 561
Qy 296 AlaHisProLysValAspProGluSer-----GlyGlu 306
Db 562 GCTCACCCCCACATTGAAATGATGGGAGCTGTTTACAACTGTTAAATGTTTGGGAAA 621
Qy 307 LeuPheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSer 326
Db 622 AATTTTCGATTCCTACAAATATTGTAAGATCCCTCCACTCCAGCAGACAAAGAGAT 681
Qy 327 ProAspGlyThrLysSerProAspValGluIleGlnLeuAsp-----Gln 341
Db 682 CCA-----ATAAGCAAGTCCGAGGTCTGCTACAAATTCCTCCGTCGAGCGACCGATTCAAG 735
Qy 342 ProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGlnGln 361
Db 736 CCATCGTACGTCATAGTTTGGTTGTGACTCCCACTATATTGTTTGTGGAGACGCCA 795
Qy 362 ValValPheLysLeuProGluMetIleArgGlyGlySer----- 374
```

```
Db 796 GTCAAAATTAACCTGCTCAAGTTCTTTCTTCGTGGAGTCTTTGGGAGGCAACTACATG 855
Qy 375 -----ProValValTyrAspLysAsnLysValAla 384
Db 856 GATTGTTTGAAGTCCCAATGAACATGGGGGTTTGGCTTCACATCGCTGACAAAAAAGA 915
Qy 385 ArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLysTyrIleAspAla 404
Db 916 AAAAAGTATCTCAATAATAAGTACAGGACCTCTTCCTTTAATCTC----- 960
Qy 405 ProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluProGluThrAspGluVal 424
Db 961 -----TTCATCATATCAATACATCTACGAGACAATGAGTTT-----CTG 999
Qy 425 ValValIleGlySerCysMetThrProProAspSerIlePheAsnGluSerAsp---Glu 443
Db 1000 ATTGTGGATCTCTGCTGCTGCGAAGGATTTGAATTCGTCTACAATTAATCTTTAGCC 1059
Qy 444 AsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGluSerThrArg 463
Db 1060 AATTACCTGAGAACTGGGAGAGGTGAAA-----AAAAATGCCAGAAAGGCTCCG 1110
Qy 464 ArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMetValAsnArg 483
Db 1111 CAGCCTCAAGTTAGGAGATCCGTCTTCCTTTGAATATCGACAAGCGCGACACAGGCAAG 1170
Qy 484 AsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeu----- 497
Db 1171 AACCTAGTCACTCCCTTCCCAACACGCGCCACTGCAACTCTGCGCAGCAGACGACCATC 1230
Qy 498 ---AlaGluProTrpProLysValSerGly-----PheAlaLysVal 510
Db 1231 TGGCTGGAACCTGAGGTTCTCTCTCAGGCGCTCGTCAAGCCTTTGAGTTTCTCAATC 1290
Qy 511 AspLeuThrThr-----GlyGluValLysLysHisLeuTyrGly----- 523
Db 1291 AACTATCAGAAGTATGCGCGGAAGCCTTACACGTACGCTGACGTTGAGCTTGGCTTGAATCAC 1350
Qy 523 ----- 523
Db 1351 TTCGTTCCGACAGGCTCTGCAAGCTGAAGCTCAAGACTAAAGAAACGTTGGTATGCAG 1410
Qy 524 ---AspAsnArgTyrGlyGlyProLeuPheLeuPro---GlyGluGlyGlyGluGlu 541
Db 1411 GAGCCCGACTCATACCCATCAGAACCCATCTTTGTTCTCACCAGATGCTTGGAGAA 1470
Qy 542 AspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLys-----SerGlu 559
Db 1471 GATGATGCTGTAGTTCTGAGTGTGTGAGTGGGCGCTGGGCGAGGACAAAGCCCTGCTTAT 1530
Qy 560 LeuGlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArg 579
Db 1531 CTTCTGATTCTGAATGCCAAGGATTTGAGTGAAGTTCCAGGCGTGAAGTGGAGATTAAAC 1590
Qy 580 ValProTyrGlyPheHisGlyThrPhe 588
Db 1591 ATCCCTGTCACCTTTCATGAGCTGTTT 1617

RESULT 11
US-08-311-731A-140
; Sequence 140, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
```

```

, STATE: MASSACHUSETTS
, COUNTRY: USA
, ZIP: 02210
, COMPUTER READABLE FORM:
, MEDIUM TYPE: Floppy disk
, COMPUTER: IBM PC compatible
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: PatentIn Release #1.0, Version #1.25
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/311,731A
, FILING DATE:
, CLASSIFICATION: 530
, ATTORNEY/AGENT INFORMATION:
, NAME: GATES, EDWARD R.
, REGISTRATION NUMBER: 31,616
, REFERENCE/DOCKET NUMBER: C0044/7125
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: 617/720-3500
, TELEFAX: 617/720-2441
, INFORMATION FOR SEQ ID NO: 140:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 36063 base pairs
, TYPE: nucleic acid
, STRANDEDNESS: double
, TOPOLOGY: circular
, MOLECULE TYPE: DNA (genomic)
, HYPOTHETICAL: NO
, ANTI-SENSE: NO
, ORIGINAL SOURCE:
, ORGANISM: Mycobacterium leprae
, US-08-311-731A-140

```

Alignment Scores:	
Pred. No.:	2.74e-08
Score:	186.00
Percent Similarity:	38.61%
Best Local Similarity:	25.05%
Query Match:	5.90%
DB:	4
Gaps:	23
Indels:	121
Mismatches:	209
Conservative:	72
Matches:	133
Length:	36083

US-09-758-269-6 (1-599) x US-08-311-731A-140 (1-36063)

Qy	110	LeuHisProLeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPhaAlaPro	129
Db	22576	TTGTCAAGCCTATAGAGGGATCACGACATGGAGTGTGCGAAATGTGCAGCAAGTCTCTGTCT	22635
Qy	130	Val-----AsnGlnProValArgArgAsnLeuPro-----	140
Db	22636	ACCTGTCCGAAGACGACAGCCATCTGCGCCCGGATGTCCCGTGGCGACCGAGACGATC	22695
Qy	141	-----ValValGlyLysLeuProAspSerIleLysGly	151
Db	22696	GAATGGAGCCCAACGACCTCAACGCTTTAGCGGGAGAAATACCCAGCTACTATAAAGCG	22755
Qy	152	valtyrValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyVHisPhePhe	171
Db	22756	TACCTACCGACGCAACACTTAAAAACCATCGACCCCGAGTTTCACAGTGTATCATCCCGCTTC	22815
Qy	172	AspGlyAspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyrAlaCys	191
Db	22816	GATGAAGACGGCATGCTGCATGTCGCTTCGCGATGGAGAAAGCCCTTTTTCGCAAC	22875
Qy	192	ArgPheThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPhePro	211
Db	22876	AGCTTTATACGACACAGACGGATTTTGTCCGAGAACATCGCAGGCGAGCCGCTGTGGGCC	22935
Qy	212	LysAlaIleGlyGluLeuHisGlyHisThrGlyVlleAlaArgLeuMetPheTyrAla	231
Db	22936	GG-TTC-----GGCAGAACCGT-----	22952
Qy	232	ArgAlaAlaAlcGlyIleValAspProAla---HisGlyThrGlyValAlaAsnAlaGly	250
Db	22953	-----GCAATATTCACCAAGCGTGAACAGGCTGGGATCCGCAACAGCATG	23000

Qy 565 AlaValSerLeuGluValGluAlaThr 573  
Db 23982 CCGTATTCAGCAGCGCGCTTCCACC 24008

RESULT 12  
US-09-252-991A-7056/c  
; Sequence 7056, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7056  
; LENGTH: 4242  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7056

Alignment Scores:  
Pred. No.: 0.00236 Length: 4242  
Score: 127.50 Matches: 104  
Percent Similarity: 35.96% Conservative: 65  
Best Local Similarity: 22.13% Mismatches: 174  
Query Match: 4.05% Indels: 129  
DB: 4 Gaps: 21

US-09-758-269-6 (1-599) x US-09-252-991A-7056 (1-4242)

Qy 23 ProLeuSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMetAlaSer 42  
Db 2486 CCATGCACTGCCGCGGTATCGAGTCTTGGCATGC-----GCCGCG 2442

Qy 43 ArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProProAlaLeuHisPhe 62  
Db 2441 AGGATTGGCACTGCACCTTCTATTCTGATCACTTCGACGCTGCCAGCGCGCTGTGG 2382

Qy 63 ProLysGlnSerSerAsnSerProAlaIleValLysProLysAlaLysGluSerAsn 82  
Db 2381 CCCCCTTGTCTGTGC--GCGCCCGCGTGTGTCGCGGCCCGCAGGCGCAGTGGGCGCGG 2325

Qy 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAlaGluGly 102  
Db 2324 AGGAAA-----TCTGGGAG-CTGATCCGCGCGGAGGCG 2293

Qy 103 PheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerValGln 122  
Db 2292 -GTGAG-CATCTCTCGGCTTACCCCGAGTACGG----- 2261

Qy 123 IleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuProValVal 142  
Db 2260 -----CAGCCAGCTCGC----- 2240

Qy 143 GlyLysLeuProAspSerIleLysGlyValTyrValArgAsn-GlyAlaAsnProLeuHi 162  
Db 2239 GGAGGCCAGCGCGCGGATGTCGCGGTGCGATGTCATCATCCGCGCGC----- 2191

Qy 162 sGluProValThrGlyHisHis-----ph 170  
Db 2190 -GAAGCGCTACCGCGAGACCTGCAACGATTCCGACGCGCTTGGCCCGCGGTGCTT 2132

Qy 170 ePheAspGlyAspGlyMetValHisAlaVal----- 180  
Db 2131 CTTCAATGCCCTACGACCGCAGCGGTGGTGCATGCCGCTGCGCTGCTGCCG 2072

Qy 181 -LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVa 200

Db 2071 ACGACTCGAGAGGGTCCGCCAGCGTCCGATCGGCAGCGTGTGTCCGCGCGCGGTTCGC 2012  
Qy 200 lGlnGluArgGlnLeuGlyArgProValPheProLys--AlaIleGlyGluLeuHis-- 218  
Db 2011 CTACATCTCGATGCCACCTGGCCCTGTGGTCCGCGAGCGGCGCGCACTTCTACGT 1952  
Qy 219 -GlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaGly----- 236  
Db 1951 CGGCGCGCGCGCTGCGCGC-----GGTACCATGAGCGTCCGCGCTCAGCGCGCA 1898  
Qy 237 -----IleValAspProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPh 254  
Db 1897 GCGCTTCGTGCCCATCCCTTC-----GCTGCCGAGGCGCGCGCTGATCCG 1850  
Qy 254 eAsnGlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIleThrPr 274  
Db 1849 CACGCGGACCTG-----GTGGCGCTGTGCGA 1823  
Qy 274 oAsnGlyAspLeuLysThrValGlyArgPheAspPheAse----- 287  
Db 1822 CACGCGCAGGTGAATATGTCCGCCCATCGACACCGAGTGAAGATCCGTGCTTCCG 1763  
Qy 288 -----GlyGlnLeuGluSerThrMetIleAlaHisProLysValAspProGluSe 304  
Db 1762 CATCGAACTGGCGAGATCGAGCGCGCTGTGGAGCATCCGCGAGTT----- 1714  
Qy 304 rGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheAr 324  
Db 1713 -CGCAAGCGCTGTGCTCGCTCGACAGCCCGAGCGCAAGCAGTTGGCGGTTATGT 1655  
Qy 324 gPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAspGlnProThrMe 344  
Db 1654 CGCCAGCGCGTGGCGCAGCAGCAGCGCGCTGCGCGCGCTGCGCGAGCGCTGAA 1595  
Qy 344 tMetHisAspPheAlaIleThrGluAsnPheValValProAspGlnGlnValAlPh 364  
Db 1594 AACGCATCTCAAGCAGCAATTGCCGACTACATGTCGCGGCCACCTGTGTTGCTCGC 1535  
Qy 364 eLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLysAsnLysValAl 384  
Db 1534 CAGCTGCGC-----CTGAC 1520  
Qy 384 aArgPheGlyIleLeuAspLysTyrAla-----GluAspSerSerAsnIleLy 400  
Db 1519 CGCCAAACGGAAGCTCAGCCGCGCGCTGCCGCGCCCGACCGCGCTCAACCGCA 1460  
Qy 400 sTrpIleAspAlaPro---AspCysPheCysPheHisLeuTrpAsnAlaTrpGluPr 419  
Db 1459 GGCCTACAGGCCCCCGCAGCGTGTGGAGCAGCAACTGCGCGGGGTCTGGCGGAGGT 1400  
Qy 419 oGluThrAspGluValValIleGlySerCysMetThrProAspSerIlePheAs 439  
Db 1399 GCTGAACGTGAGCGGTAGGTCTCGGC-----GACAACTTCTTCGA 1358  
Qy 439 nGluSerAspGluAsnLeuLysSerVal 448  
Db 1357 ACTGGCGCGCATTCGATCTGTGATC 1330

RESULT 13  
US-09-252-991A-6997  
; Sequence 6997, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190

;; PRIOR FILING DATE: 1998-07-27

;; NUMBER OF SEQ ID NOS: 33142

;; SEQ ID NO 6997

;; LENGTH: 10023

;; TYPE: DNA

;; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-6997

# Alignment Scores:

Pred. No.: 0.0102 Length: 10023  
 Score: 127.50 Matches: 104  
 Percent Similarity: 35.96% Conservative: 65  
 Best Local Similarity: 22.13% Mismatches: 174  
 Query Match: 4.05% Indels: 129  
 DB: 4 Gaps: 21

US-09-758-269-6 (1-599) x US-09-252-991A-6997 (1-10023)

QY 23 ProLeuSerSerSerClnSerSerAspLeuSerTyrCysSerSerLeuProMetAlaSer 42  
 Db 6671 CCATGCACCTGCGCGCGGTGATCGAGTGTCTCGCATGC-----GCGCG 6715  
 QY 43 ArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProProAlaLeuHisPhe 62  
 Db 6716 AGATTGCGAATGCACTTCTATTGATCACTTCGACCTGCCAGCGAGCGCTGCTGG 6775  
 QY 63 ProLysGlnSerSerAsnSerProAlaIleValLysProLysAlaLysGluSerAsn 82  
 Db 6776 CCCCCTGTGTGTGCG---CGCGCGCGTGTGTCGCGGCCCGAGGCCAGTGGCGCGG 6832  
 QY 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaGluGly 102  
 Db 6833 AGGAAA-----TCTCGAG-CTGATCGCGCGAGGCGC 6864  
 QY 103 PheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerValGln 122  
 Db 6865 -----GTGAG-CATCTCGGCTTACCCCGAGCTACGG----- 6896  
 QY 123 IleAlaGlyAsnPheAlaProValAsnGluGlnProValArgAsnLeuProValVal 142  
 Db 6897 -----CAGCCAGCTCGC-----CCAGTGGCT 6917  
 QY 143 GlyLysLeuProAspSerIleLysGlyValTyrValArgAsn-GlyAlaAsnProLeuHi 162  
 Db 6918 GGAGAGCCAGGCGCGAGTTCGCGTGCATGTGTCATCACCAGCGCGC----- 6966  
 QY 162 sGluProValThrGlyHisHis-----Ph 170  
 Db 6967 -GAAGCGCTGACCGCGGACCTGCAACGGATTGCCAGGCTTCGCCCGCGCGCTGT 7025  
 QY 170 ePheAspGlyAspGlyMetValHisAlaVal----- 180  
 Db 7026 CTTCAATGCTACGAGACCGAGAGCGGTGTGTCGCGTGGCTGCTGCTGCCCGA 7085  
 QY 181 -LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVa 200  
 Db 7086 ACCACTGGAGGAGGTGCGCGCAGCGTGCAGTGCAGCGTGTGCGCGCGCGGTGCG 7145  
 QY 200 lGlnGluArgGlnLeuGlyArgProValPheProLys---AlaIleGlyGluLeuHis-- 218  
 Db 7146 CTACATCTGATGCGCGACCTGCGCTGTGTCGCGAGGCGCGCGCAACTCTACGT 7205  
 QY 219 -GlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGly----- 236  
 Db 7206 CGCGCGCGCGCGCTGCGCGC-----GGCTACCATGAGCTCGCGGCTCAGCGCGCA 7259  
 QY 237 -----IleValAspProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPh 254  
 Db 7260 GCGCTTGTGTCGCGATCCCTTC-----GCTCGCGAGGCGCGCGCTGTACCG 7307  
 QY 254 eAsnGlyArgLeuLeuAlaMetSerGluAspLeuProTyrGlnValGlnIleThrPr 274  
 Db 7308 CACCGCGGACCTG-----GTGCGCGCTGTGCGA 7334

QY 274 oAsnGlyAspLeuLysThrValGlyArgPheAspPheAsp----- 287  
 Db 7335 CAACGCGCAGGTGGAATATGTGCGCGCATCGACCACCAAGATCCGTGGCTTCG 7394  
 QY 288 -----GlyGlnLeuGluSerThrMetIleAlaHisProLysValAspProGluSe 304  
 Db 7395 CATCGAACTGGCGGAGATCGAGCGCGCTGTGTGAGCATCCGCAAGTT----- 7443  
 QY 304 rGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheAr 324  
 Db 7444 -CGCAAGCGCTGCTGCTGCGCTCGACAGCCCGAGCGCAAGTGGCCGGTTATGT 7502  
 QY 324 gPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAspGlnProThrMe 344  
 Db 7503 CGCCAGCGCGTGGCGGAGCAGGACGCGCCAGCGCGCTGCGCGAGCGGTGAA 7562  
 QY 344 tMetHisaspPheAlaIleThrGluAsnPheValValProAspGlnValValPh 364  
 Db 7563 AACGCATCTCAAGCAGCAATTGCGGACTACATGTGCGGCGCCACCTGCTGCTCGC 7622  
 QY 364 eLysLeuProGluMetIleArgGlySerProValValTyrAspLysAsnLysValAl 384  
 Db 7623 CAGCTGCGC-----CTGAC 7637  
 QY 384 aArgPheGlyIleLeuAspLysTyrAla-----GluAspSerSerAsnIlely 400  
 Db 7638 CGCCAAACGCAAGCTCGACCGCGCGCTGCGCGCGCGCGCTCAACCGGCA 7697  
 QY 400 sTrpIleAspAlaPro---AspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluPr 419  
 Db 7698 GGCCTACAGCGCGCGCGCGCGCTGCTCGAGCAGCAACTGCGCGCGGTCTGCGCGAGGT 7757  
 QY 419 oGluThrAspGluValValValIleGlySerCysMetThrProProAspSerIlePheAs 439  
 Db 7758 GCTGACCTCAGCGGGTAGGTCTCGGC-----GACAACTTCTTGA 7799  
 QY 439 nGluSerAspGluAsnLeuLysSerVal 448  
 Db 7800 ACTGGCGCGCATTCGATCTCTCGATC 7827

## RESULT 14

US-08-700-651-1  
 ; Sequence 1, Application US/08700651B  
 ; Patent No. 6015882

### GENERAL INFORMATION:

;; APPLICANT: PETERSEN, CAROLYN  
 ;; APPLICANT: LEECH, JAMES  
 ;; APPLICANT: NELSON, RICHARD, C.  
 ;; APPLICANT: GUT, JIRI  
 ;; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS  
 ;; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum  
 ;; TITLE OF INVENTION: INFECTIONS  
 ;; FILE REFERENCE: 480.19-4 (HV)  
 ;; CURRENT APPLICATION NUMBER: US/08/700.651B  
 ;; CURRENT FILING DATE: 1997-08-14  
 ;; EARLIER APPLICATION NUMBER: 08/415,751  
 ;; EARLIER FILING DATE: 1995-04-03  
 ;; NUMBER OF SEQ ID NOS: 15  
 ;; SOFTWARE: PatentIn ver. 2.0  
 ;; SEQ ID NO 1  
 ;; LENGTH: 5163  
 ;; TYPE: DNA  
 ;; ORGANISM: Cryptosporidium parvum  
 US-08-700-651-1

### Alignment Scores:

Pred. No.: 0.0092 Length: 5163  
 Score: 123.50 Matches: 125  
 Percent Similarity: 34.60% Conservative: 85  
 Best Local Similarity: 20.59% Mismatches: 194  
 Query Match: 3.92% Indels: 203  
 DB: 3 Gaps: 35

```
US-09-758-269-6 (1-599) x US-08-700-651-1 (1-5163)
QY 5 ThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThrGlnProLeu 24
Db 2491 ACAATTGCAGGTATTGTTTCAGGA-----ATT 2517
QY 25 SerSerGlnSerSerAspLeuSerTyrCysSerSerLeu-----ProMetAlaSer 42
Db 2518 TCTGCAAGTGAGTCAATTA---TTATCTCAGAAATCAGCTCTAATCGACCCAGCAACAAAT 2574
QY 43 ArgValThrArgLys-----LeuAsnValSerSerAlaLeuHisThrProPro 58
Db 2575 ATGTTGTTGGAGAAATTTGGTGGATTGTTGAACCCAGCAACAGGAGTGATGATCCAGGT 2634
QY 59 AlaLeuHisPheProLysGlnSerSerAsnSerProAlaIle-----ValVal 74
Db 2635 TTTTGTAGTCCATCAGAGCAAACTCAATCTCCCTGAGATTGAAGATTGAGTGGTATTATT 2694
QY 75 LysProLysAlaLysGluSerAsnThrLysGlnMetAsnLeu-----88
Db 2695 CCTCCAGAAAGTAGCAGCAGCAAACTGCTGATAAATTCAGATTATCTATTCCTCCAGCGTA 2754
QY 89 -----PheGlnArgAlaAlaAlaAlaAlaLeuAsp 98
Db 2755 CCAGAAATCAATTCAGAAAAGGATCAGAAAGATTGATTCTATTCTGAATTGATGATGAT 2814
QY 99 AlaAlaGluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAsp 118
Db 2815 ATTGAGTCAGGTAGACTTATTGGTCAAGTATCAAGAGACCAATCCCGAGTTCA-----2868
QY 119 ProSerValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsn 138
Db 2869 -----ATTGCTGGTGACTTGAACCAATAATGAAGACACCAACAACACTGAC 2916
QY 139 LeuProValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAla 158
Db 2917 ---AGTGAATGCTGTTAAACCAATCGATCCACACAGGTCTG-----2955
QY 159 AsnProLeuHisGluProValThrGlyHisHisPhePheAspGlyAspGlyMetValHis 178
Db 2956 ---CCTTTCAATCCACCA---ACTGGTCAAT-----TTGATTAAC 2988
QY 179 AlaValLysPheGluHisGlySerAlaSerTyr-----AlaCysArgPheThrGlnThr 196
Db 2989 CCAACAAATAATAATACCATGGATTCCTTCATTGCTGTGTCATCAAAATATGCGAGTTTCA 3048
QY 197 AsnArgPheValGlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGlu 216
Db 3049 AATGTTATTAGACTGATAATGTTTATGTTTACCAGTT-----GGTGAA 3093
QY 217 LeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaGly 236
Db 3094 ATAACAGGTTTACCAGAAAGGATCCAGGCTCAGATATTCCATTTAACTCACTACAGGTGAA 3153
QY 237 IleValAspProAlaHisGlyThrGlyValAlaAsn-----AlaGlyLeuValTyrPhe 254
Db 3154 TTAGTTGATTCATCAACAGGAAGCAATTAACAAATTCATCTGCTGGTATTGTT-----3207
QY 255 AsnGlyArgLeu---LeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIleThr 273
Db 3208 AGTGAAAACCTGGCTTACCACCTATTGAAGTGA-----3243
QY 274 ProAsnGlyAspLeuLysThrValGlyArgPheAsp-----PheAsp 287
Db 3244 ---AATGTAATTG-----TTTGATCCATCAACTCAACTGCAATAGAT 3285
QY 288 GlyGlnLeuGluSerThrMetIleAlaHisProLysValAspProGluSerGlyGluLeu 307
Db 3286 GGTAAAT-----AACCAATTAGTTAAACCAAGAAACCAACAGCACT 3324
QY 308 PheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerPro 327
Db 3324 -----AAACCAAGTAAATGTTAAACCAAGTAAATGTTAAACCAAGTAAATGTTAAAT 3891
```

```
3325 GTCTCAGGATCAACTTCAGGTACTACAAAACCA-----AAACCA 3363
QY 328 -----AspGlyThrLysSerProAspValGluIle-----GlnLeuAsp 340
Db 3364 GGAATTCAGTCAATGTGGAGGTGTTGTCCTGATGAAGAAGCTAAAGATCAAGCCGAT 3423
QY 341 Gln-----ProThrMet 344
Db 3424 AAGGTAAGGATGATTAATTTGTTCCACCAACTAATTTCTATCAATAAAGATCCAGTAACA 3483
QY 345 MetHisAspPheAlaIleThrGluAsnPheValValProAspGlnGlnValPhe 364
Db 3484 AATACTCAGTACAGTAATACTACTGTTAAACATTATTAAACCCAGAAACAGGAAAAAGTT---3540
QY 365 LysLeuProGluMetIleArgGly-----GlySerProValVal 377
Db 3541 ---ATTCCAGGTTCCACTTCAGGCTCTCTCAACTATCCATCAATTAATCTCCACAACAA 3597
QY 378 TyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSer 397
Db 3598 ACTGATGAG-----ATTACAGAAAGCCAGTTGATGACTGTTACT 3636
QY 398 AsnIleLysTrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGlu 417
Db 3637 GGTTCGCATAT-----3648
QY 418 GluProGluThrAspGluValValIleGlySerCysMetThrProProAspSerIle 437
Db 3649 GATCCATCTACAGGTGAATTTATCGATCCTGCAACTAAATTTACCAATTTCCAGGATCAGTT 3708
QY 438 PheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLys 457
Db 3709 GCAGTGAT-----GAAATCTCTACTGAAGTA---TTGAACATACA 3747
QY 458 ThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGlu 477
Db 3748 ACAGATGAAGTAACAGGTTTGCCCAATT-----GATCTTGAA 3783
QY 478 AlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeu 497
Db 3784 ACTGGTCTT-----3792
QY 498 AlaGluProTrpProLysValSerGlyPheAlaLysVal-----AspLeuThrThr 514
Db 3793 -----CCAAGATCCAGTATCAGGATCCCACTCCCAACTCCAAATGGTACTTGGTTGAT 3846
QY 515 GlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGluProLeuPheLeu 534
Db 3847 CCATCAAAATAAAACCAATTCAGGTTCA-----CATTCGGA-----TTTATT 3891
QY 535 ProGlyGluGlyGlyGlu 541
Db 3892 AATGGTATCATCTGGGAACAA 3912
RESULT 15
US-08-928-361B-4
; Sequence 4, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA
; STREET: 365 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
```

; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/928,361B  
 ; FILING DATE: 12-SEP-1997  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 60/026,062  
 ; FILING DATE: 13-SEP-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Verny, Hana  
 ; REGISTRATION NUMBER: 30,518  
 ; REFERENCE/DOCKET NUMBER: 480.76-1 (HV)  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 650-324-1677  
 ; TELEFAX: 650-324-1678  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 5163 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; US-08-928-361B-4

Alignment Scores:  
 Pred. No.: 0.0092 Length: 5163  
 Score: 123.50 Matches: 125  
 Percent Similarity: 34.60% Conservative: 85  
 Best Local Similarity: 20.59% Mismatches: 194  
 Query Match: 3.92% Indels: 203  
 DB: 3 Gaps: 35

US-09-758-269-6 (1-599) x US-08-928-361B-4 (1-5163)

QY	159	AsnProLeuHisGluProValThrGlyHisHisPhePheAspGlyAspGlyMetValHis	178
DB	2956	---CCTTTCAATCCACCA---ACTGGTCAT---TTGATTAAC	2988
QY	179	AlaValLysPheGluHisGlySerAlaSerTyr---AlaCysArgPheThrGlnThr	196
DB	2989	CCAACAATAATAATACCATGATTCTTCATTTGGTGGTCATACAAATATGCAGTTTCA	3048
QY	197	AsnArgPheValGlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGlu	216
DB	3049	AATGGTATTAAAGACTGATAATCTTATGTTTACCAGTT---GGTGA	3093
QY	217	LeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaGly	236
DB	3094	ATAACAGGTTTACCAAGGATCCAGGCTCAGATTTTCAATTTAACTCAACTACAGGTAA	3153
QY	237	IleValAspProAlaHisGlyThrGlyValAlaAan---AlaGlyLeuValTyrPhe	254
DB	3154	TTAGTTGATCCATCAACAGGAAGCAATTAACAATTTCTACTGCTGGTATTGTT---	3207
QY	255	AsnGlyArgLeu---LeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIleThr	273
DB	3208	ACTGGAAACCTGGCTTACCACCTATTGAAGATGAA---PheAsp	287
QY	274	ProAsnGlyAspLeuLysThrValGlyArgPheAsp---PheAsp	287
DB	3244	---AATGGTAATTG---TTTGATCCATCAACTAACTTGCATATAGAT	3285
QY	288	GlyGlnLeuGluSerThrMetIleAlaHisProLysValAspProGluSerGlyGluLeu	307
DB	3286	GGTAAT---AACCAATTAGTTTAAACCAGAAACCAACAGCAGCT	3324
QY	308	PheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerPro	327
DB	3325	GTCTCAGGATCACTTCAGGTACTACAAACCA---AAACCA	3363
QY	328	---AspGlyThrLysSerProAspValGluIle---GlnLeuAsp	340
DB	3364	GAATTTCCAGTCAATGGTGGAGGTGTGTACCTGATGAAGAGTAAAGATCAAGCCGAT	3423
QY	341	Gln---ProThrMet	344
DB	3424	AAGGTAAAGGATGGATTAATTGTTCCCAACTAATTTCTATCAATAAAGATCCAGTAACA	3483
QY	345	MetHisAspPheAlaIleThrGluAsnPheValValProAspGlnGlnValPhe	364
DB	3484	ATACTCAGTACAGTAATACTACTCGTAACATTATTAACCCAGAACAGGAAAGTT---	3540
QY	365	LysLeuProGluMetIleArgGly---GlySerProValVal	377
DB	3541	---ATTCCAGGTTTCACTTCCAGGCTCTCTCAACTATCCATCATTCATCACTCCACAACA	3597
QY	378	TyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSer	397
DB	3598	ACTGATGAG---ATTACAGGAAAGCCAGTTGATCTGTTACT	3636
QY	398	AsnIleLysTyrIleAspAlaProAspCysPheHisLeuTyrAsnAlaIleTrpGlu	417
DB	3637	GGTTGGCCATAT---TTGAACATTACA	3648
QY	418	GluProGluThrAspGluValValIleGlySerCysMetThrProAspSerIle	437
DB	3649	GATCCATCTACAGGTGAAATTTATCGATCTCGCACTAAATTTACCAATTTCCAGGATCAGTT	3708
QY	438	PheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLys	457
DB	3709	CGAGTGAT---GAAATCTCTCACTGAAGTA---TTGAACATTACA	3747
QY	458	ThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGlu	477
DB	3748	ACAGATGAAGTAAACAGGTTTCCCAATT---GATCTTGAA	3783
QY	478	AlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeu	497

QY	5	ThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThrGlnProProLeu	24
DB	2491	ACAATTTGCAGGTATTGTTTCAGGA---ATT	2517
QY	25	SerSerSerGlnSerSerAspLeuSerTyrCysSerSerLeu---ProMetAlaSer	42
DB	2518	TTGCAAGTGAGTCATTA---TTATCTCAGAAATCAGCTCTAATCGACCCACCAACAAAT	2574
QY	43	ArgValThrArgLys---LeuAsnValSerSerAlaLeuHisThrProPro	58
DB	2575	ATGGTTGTTGGAGAATTTGGTGGATTGTTGAACCCAGCAACAGGAGTGATGATTCAGGT	2634
QY	59	AlaLeuHisPheProLysGlnSerSerAsnSerProAlaIle---ValVal	74
DB	2635	TTTTTAGTCCATCAGACCAACTCAATTTCCCTCGAGATTGAAGATGGTGGTATTAT	2694
QY	75	LysProLysAlaLysGluSerAsnThrLysGlnMetAsnLeu---88	88
DB	2695	CCTCCAGAAAGTAGCAGCAGCAAAATGCTGATAAAATTCAGTTATCTATCTCCTCCAAAGCGTA	2754
QY	89	---PheGlnArgAlaAlaAlaAlaLeuAsp	98
DB	2755	CCAGATCAATTTCCAGAAAGGATCAGAAGATTGATCTTCTGAATTGATGATGAT	2814
QY	99	AlaAlaGluGlyPheLeuValSerHisGluLysLeuHisProLysThrAlaAsp	118
DB	2815	ATTGAGTCAGGTAGACTTATTGGTCAAGTATCAAGAGACCAATCCAGGTTCA-----	2868
QY	119	ProSerValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsn	138
DB	2869	---ATTGCTGGTGAACCCCAATTAATGAAGACACCAACACAAACTGAC	2916
QY	139	LeuProValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAla	158
DB	2917	---AGTGTAACTGGTAAACCAATTCGATCCCAACACAGGTCG-----	2955

```

Db      3784 ACTGGTCTT----- 3792
Qy      498 AlaGluProTrpProLysValSerGlyPheAlaLysVal-----AspLeuThrThr 514
Db      3793 -----CCAAGAGATCCAGTATCAGGACTCCCAAACTTCCAAATGGGTACCTTGGTTGAT 3846
Qy      515 GlyGluValLysLysHisLeuTyxClyAspAsnArgTyrglyGlyGluProLeuPheLeu 534
Db      3847 CCATCAAAATAAAAAACCAATTCCAGGTTCA-----CATTCCGGA-----TTTATT 3891
Qy      535 ProGlyGluGlyGlyGluGlu 541
Db      3892 AATGGTACATCTGGGAGAACAA 3912

```

Search completed: November 15, 2003, 20:02:48  
 Job time : 14378 secs



GenCore version 5.1.6  
-Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 15, 2003, 15:07:29 ; Search time 535 Seconds  
(without alignments)  
3659.167 Million cell updates/sec

Title: US-09-758-269-6

Perfect score: 3150

Sequence: 1 MASFTATAVSCRWLGNNHT.....VPYGFHGTGIGADDLAKQVV 599

Scoring table:

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09758269/runat\_14112003\_135800\_28838/app.query.fasta\_1.775  
-DB=Published Applications NA -QFWT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MAIRX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09758269 @CIGN 1 1 347 @runat\_14112003\_135800\_28838  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/1/pubnpa/US07\_PUBCOMB.seq:
- 2: /cgn2\_6/ptodata/1/pubnpa/PCT\_NEW\_PUB.seq:
- 3: /cgn2\_6/ptodata/1/pubnpa/US05\_NEW\_PUB.seq:
- 4: /cgn2\_6/ptodata/1/pubnpa/US06\_PUBCOMB.seq:
- 5: /cgn2\_6/ptodata/1/pubnpa/US07\_NEW\_PUB.seq:
- 6: /cgn2\_6/ptodata/1/pubnpa/US08\_PUBCOMB.seq:
- 7: /cgn2\_6/ptodata/1/pubnpa/US09\_NEW\_PUB.seq:
- 8: /cgn2\_6/ptodata/1/pubnpa/US08\_PUBCOMB.seq:
- 9: /cgn2\_6/ptodata/1/pubnpa/US09A\_PUBCOMB.seq:
- 10: /cgn2\_6/ptodata/1/pubnpa/US09B\_PUBCOMB.seq:
- 11: /cgn2\_6/ptodata/1/pubnpa/US09C\_PUBCOMB.seq:
- 12: /cgn2\_6/ptodata/1/pubnpa/US09\_NEW\_PUB.seq:
- 13: /cgn2\_6/ptodata/1/pubnpa/US10A\_PUBCOMB.seq:
- 14: /cgn2\_6/ptodata/1/pubnpa/US10B\_PUBCOMB.seq:
- 15: /cgn2\_6/ptodata/1/pubnpa/US10\_NEW\_PUB.seq:
- 16: /cgn2\_6/ptodata/1/pubnpa/US60\_NEW\_PUB.seq:
- 17: /cgn2\_6/ptodata/1/pubnpa/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3150	100.0	1800	10	US-09-758-269-5 Sequence 5, Appli

2	2280.5	72.4	1818	10	US-09-758-269-15	Sequence 15, Appli
3	2168.5	68.8	1839	10	US-09-758-269-11	Sequence 11, Appli
4	1991	63.2	1752	10	US-09-758-269-1	Sequence 1, Appli
5	1930	61.3	1815	10	US-09-758-269-13	Sequence 13, Appli
6	1663.5	52.8	1734	10	US-09-758-269-9	Sequence 9, Appli
7	969	30.8	1788	10	US-09-758-269-3	Sequence 3, Appli
8	969	30.8	1788	10	US-09-938-842A-1444	Sequence 1444, Ap
9	938	29.8	1617	10	US-09-758-269-7	Sequence 7, Appli
10	937	29.7	1617	10	US-09-758-269-17	Sequence 17, Appli
11	522	16.6	393	10	US-09-878-574-2543	Sequence 2543, Ap
C 12	419	13.3	32679	8	US-08-976-063C-1	Sequence 1, Appli
C 13	411.5	13.1	1518	8	US-08-976-063C-21	Sequence 21, Appli
14	369	11.7	320	10	US-09-878-574-2872	Sequence 2872, Ap
C 15	346	11.0	200	9	US-09-770-696-172	Sequence 172, App
16	279	8.9	534	12	US-10-149-759-65	Sequence 65, Appli
17	260	8.3	3111	14	US-10-053-192-2	Sequence 2, Appli
18	238.5	7.6	2134	12	US-10-168-517-18	Sequence 18, Appli
19	234.5	7.4	367	9	US-09-770-791-402	Sequence 402, Appli
20	232.5	7.4	1855	12	US-10-168-517-16	Sequence 16, Appli
21	207.5	6.6	1934	12	US-10-168-517-20	Sequence 20, Appli
22	150	4.8	2037	12	US-10-168-517-1	Sequence 1, Appli
23	133	4.2	82993	15	US-10-080-170-645	Sequence 645, App
24	127.5	4.0	13029	9	US-09-815-242-4052	Sequence 4052, Ap
25	122	3.9	5127	14	US-10-156-761-3630	Sequence 3630, Ap
C 26	122	3.9	902608	14	US-10-156-761-1	Sequence 1, Appli
27	111.5	3.5	1875	14	US-10-156-761-3143	Sequence 3143, Ap
28	111	3.5	1956	14	US-10-156-761-4583	Sequence 4583, Ap
29	110	3.5	902608	14	US-10-156-761-1	Sequence 1, Appli
30	108.5	3.4	4146	12	US-10-032-585-6784	Sequence 6784, Ap
31	107	3.4	2825	12	US-10-289-757-15	Sequence 15, Appli
32	107	3.4	3075	9	US-09-833-790-424	Sequence 424, App
33	107	3.4	7515	12	US-10-101-510-202	Sequence 202, App
34	107	3.4	7515	12	US-10-210-120-4	Sequence 4, Appli
35	107	3.4	81940	10	US-09-759-508B-1	Sequence 1, Appli
36	107	3.4	81940	12	US-09-873-319-724	Sequence 724, App
37	107	3.4	81940	12	US-09-960-706-1092	Sequence 1092, Ap
C 38	106.5	3.4	1083	9	US-09-815-242-7996	Sequence 7996, Ap
39	106.5	3.4	10174	14	US-10-171-311-82	Sequence 82, Appli
40	106	3.4	1239	10	US-09-922-261-76	Sequence 76, Appli
41	106	3.4	1825	10	US-09-922-261-75	Sequence 75, Appli
42	106	3.4	2514	14	US-10-156-761-2792	Sequence 2792, Ap
43	106	3.4	4308	12	US-10-102-549-1	Sequence 1, Appli
44	106	3.4	4331	12	US-10-354-358-67	Sequence 67, Appli
45	105.5	3.3	3048	14	US-10-156-761-3146	Sequence 3146, Ap

ALIGNMENTS

RESULT 1  
US-09-758-269-5  
; Sequence 5, Application US/09758269  
; Patent No. US20020104120A1  
; GENERAL INFORMATION:  
; APPLICANT: IUCHI, SATOSHI  
; APPLICANT: KOBAYASHI, MASATOMO  
; APPLICANT: SHINOZAKI, KAZUO  
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN  
; TITLE OF INVENTION: CLEAVAGE ENZYME GENE  
; FILE REFERENCE: 3914-3  
; CURRENT APPLICATION NUMBER: US/09/758,269  
; CURRENT FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: JP 2001-003476  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: JP 2000-010056  
; PRIOR FILING DATE: 2000-01-13  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 1800  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: CDS

i LOCATION: (1)...(1797)  
US-09-758-269-5

## Alignment Scores:

Pred. No.: 0 Length: 1800  
Score: 3150.00 Matches: 599  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-758-269-6 (1-599) x US-09-758-269-5 (1-1800)

```
Qy 1 MetAlaSerPheThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThr 20
Db 1 ATGGCTTCTTTCACGGCAACGGCTGCGGTTTCTGGGAGATGGCTTGGTGGCAATCATACT 60

Qy 21 GlnProProLeuSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet 40
Db 61 CAGCGCCATATATCGTCTTCTCAAGCTCCGACTTGAGTTATTTAGTCTCTTACCTATG 120

Qy 41 AlaSerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProProAlaLeu 60
Db 121 GCCAGTCGTGCACACGTAAGCTCAATGTTTCATCTGGCTTCACACTCCTCCAGCTCTT 180

Qy 61 HisPheProLysGlnSerSerAsnSerProAlaIleValValLysProLysAlaLysGlu 80
Db 181 CATTTCCCTAAGCAATCATCAAACTCTCCGCCATTTGTTAAGCCCAAGCCAAAGAA 240

Qy 81 SerAsnThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAla 100
Db 241 TCCACACTAAACAGATGAATTTGTTCCAGAGCGCGCGCGGAGCGTTTGAGCGCGCG 300

Qy 101 GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer 120
Db 301 GAGGTTTCTTGTGAGCCACGAGAAGCTACACCGCTTCTCAAAACGGCTGATCTTAGT 360

Qy 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuPro 140
Db 361 GTTCAGATCGCGCGGAAATTTTGTCTCCGGTGAATGAACACCGCGCTCGCGGTAATCTCCG 420

Qy 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160
Db 421 GTGGTCGGAACACTTCCCGATTCATCAAGAGAGTGATATGTGCGCAACGGAGCTTAACCCA 480

Qy 161 LeuHisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaVal 180
Db 481 CTTCAAGAGCGGTGACAGGTCAACCACTTCTTCGACGGAGACGGTATGGTTTACGCGGTC 540

Qy 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200
Db 541 AAATTCGAACACGGTTACGTAGCTAGCTTACCGCTTTCGCCGTTTACTCAGACTAAACCGGTTTGT 600

Qy 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220
Db 601 CAGGAACGTCAATGGGTGCGACCGGTTTCCCAAGCCATCGGTGAGCTTCACGGCCAC 660

Qy 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspPro 240
Db 661 ACCGGTATTCGCCGACTCATGCTATTCACGCCAGAGCTGCAGCGGTATAGTCGACCCG 720

Qy 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260
Db 721 GCACACGGAACCGGTGTAGCTAACCGCGTTTGGTCTATTTCAATGGCGCGTTTATGGCT 780

Qy 261 MetSerGluAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280
Db 781 ATGTCGGAGGATGATTACCTTACCAAGTTTCCAGTCACTCCCAATGGAGATTTAAACACC 840

Qy 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300
Db 841 GTTGTGCGGTTCGATTTTGTATGGACAAATTAGATCCCAATGATTTGCCCAACCCGAAAGTC 900
```

```
Qy 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeu 320
Db 901 GACCCGGAATCCGGTGAACCTTTCGCTTTAAGCTACGACGTCGTTTCAAAGCCCTTACCCTA 960

Qy 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340
Db 961 AAATACCTCCGATTTCTCACCAGGAGAACTAAATCACCAGGACGTCGAGATTCAGCTTGAT 1020

Qy 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGln 360
Db 1021 CAGCAACGATGATGCACGATTTCCGATTTACAGAACTTCGTCGTCGTACCTGACCCAG 1080

Qy 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380
Db 1081 CAAGTCGTTTTCAGCTGCGGAGATGATCGCGGTGGGTCTCCGGTGGTTTACGACAAG 1140

Qy 381 AsnLysValAlaAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400
Db 1141 AACAAAGTCGCAAGATTCGGGATTTTAGACAAATACGCCGAAGATTCATCGAACATTAAG 1200

Qy 401 TrpIleAspAlaProAspCysPhePheHisLeuTrpAsnAlaTrpGluGluProGlu 420
Db 1201 TGGATTGATGCTCCAGATTTGCTTCTGCTTCATCTCTGGAACGCTTGGGAAGACCCAGAA 1260

Qy 421 ThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440
Db 1261 ACAGTGAAGTCGTCGTGATAGGTCCTGTATGACTCCACCAGACTCAATTTTCAACGAG 1320

Qy 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460
Db 1321 TCTGACGAGAATCTCAAGAGTGCTCTGCTGAAATCCGCTGAAATCTCAAAACCGGTGAA 1380

Qy 461 SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480
Db 1381 TCAACTCGCGTCGATCATCTCCCAAGAGATCAACAAGTCAACCTCGAAGCAGGGATG 1440

Qy 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500
Db 1441 GTCAACAGAAACATGCTCGCGCGTAAACCAAAATTCGCTTACTTTGGCTTTAGCCGAGCGG 1500

Qy 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis 520
Db 1501 TGGCTTAAGTCTCAGGATTCGCTAAAGTTGATCTCTACTCGAGAAGTTTAAGAACAT 1560

Qy 521 LeuTyrGlyAspAsnArgTyrGlyGlyGluProLeuPheLeuProGlyGluGlyGlu 540
Db 1561 CTTTACGCGATAACCGTTACGGAGGAGAGCCTCTGTTTCTCCCGGAGAAGGAGAGAG 1620

Qy 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560
Db 1621 GAAGCAAGAGATACCTCTCTGTTTCGTTCCAGCAGAGAAGACATGGAAATCCGAGTTA 1680

Qy 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580
Db 1681 CAGATAGTTAAACCGCTAGCTTAGAGTTGAACCAACGGTTAAACTTCCTCGTCAAGGGTT 1740

Qy 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspLeuAlaLysGlnValVal 599
Db 1741 CCGTACGGAATTCACGGTACATTCATCGAGCCGATGATTTGGCGAAGCAGGTCGTG 1797
```

## RESULT 2

US-09-758-269-15  
; Sequence 15, Application US/09758269  
; Patent No. US20020104120A1  
; GENERAL INFORMATION:  
; APPLICANT: IUCHI, SATOSHI  
; APPLICANT: KOBAYASHI, MASATOMO  
; APPLICANT: SHINOZAKI, KAZUO  
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN  
; FILE REFERENCE: 3914-3  
; CURRENT APPLICATION NUMBER: US/09/758,269  
; CURRENT FILING DATE: 2001-01-12

```
; PRIOR APPLICATION NUMBER: JP 2001-003476
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: JP 2000-010056
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1815)
US-09-758-269-15

Alignment Scores:
Pred. No.:      4,34e-252      Length:      1818
Score:          2280.50        Matches:     435
Percent Similarity: 82.04%    Conservative: 63
Best Local Similarity: 71.66% Mismatches:     90
Query Match:       72.40%     Indels:         19
DB:                10         Gaps:           6

US-09-758-269-6 (1-599) x US-09-758-269-15 (1-1818)

Qy   5 ThrAlaThrAlaValSerGlyArgTrpLeuGlyGlyAsnHisThrGlnProProLeu 24
Db   7 ACTACTACTTCACATGCCCAATAACATCGATT-----AAGACTAAAGTTG 51
Qy   25 SerSerSerGlnSerSerAspLeuSerTyrCysSer-----SerLeuProMetAla 41
Db   52 TCAATGCCATCATCAAGAGAGTTGGTTTTGCATCAAACCTCTATTCTCTACTCAAAAAT 111
Qy   42 SerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProAlaLeuHis 61
Db   112 CAACATAAATAGGCCAAAGTCTCAACATTAATTCCTCTCTTCAAGCTCCACCTATACTTCAT 171
Qy   62 PheProLysGlnSerSerAsn-----SerProAla-----IleValValLysProLys 77
Db   172 TTTCCTTAAACAATCTTTCAAATATTCAAACACCACCAAGAATAATACAAATTTCAACCCAAAA 231
Qy   78 AlaLysGluSerAsn-----ThrLysGlnMetAsnLeuPheGlnArgAla 92
Db   232 CAGAAACAACAACACTCTCTCTCTTCTCAACTCCAAGTGGAATTTAGTCGAGAAAGCA 291
Qy   93 AlaAlaAlaLeuAspAlaAlaGluGlyPheLeuValSerHisGluLysLeuHisPro 112
Db   292 GCAGCAATGGCTTTAGATGCTGTGAGAAAGTGCTTTTAACATAAACATGAACCTTGAACACCCCT 351
Qy   113 LeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValAsnGlu 132
Db   352 TTCGCCGAAACAGCCGACCACCGAGTCGACAGTTCTCGGGAATTTTGCTCGGGTACCGGAA 411
Qy   133 GlnProValArgArgAsnLeuProValValGlyLysLeuProAspSerIleLysGlyVal 152
Db   412 AATCCAGTCTGTCAATCTCTTCGGGTACC GGAAAAAATACCCAATGTGTTCAGGGGTT 471
Qy   153 TyrValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHisHisPhePheAsp 172
Db   472 TAGCTTCGAAACGGAGCTTAACCTCTTTTTTGAACCAACCCGCGGACACCACTTTCTTCGAC 531
Qy   173 GlyAspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyrAlaCysArg 192
Db   532 GCGCAGCGGTATGGTTCACGCCGTTCAATTCAAAAATGGGTGGGTAGTAGTTCAGCTTGCCT 591
Qy   193 PheThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPheProLys 212
Db   592 TTCCTGAAACAGAGAGGCTGTTTCAGAAAAAGCTTTGGGTCCGCTGTTTCCCTTAA 651
Qy   213 AlalleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArg 232
Db   652 GCCATTGGTCAATTACATGGTCACCTCGAATTCGAAGCTTATGCTGTTTACGCTCGT 711
```

```
Db      1792 GCCAATGATTGGCAAAATCAG 1812
||||:|||||
RESULT 3
US-09-758-269-11
; Sequence 11, Application US/09758269
; Patent No. US20020104120A1
; GENERAL INFORMATION:
; APPLICANT: IUCHI, SATOSHI
; APPLICANT: KOBAYASHI, MASATOMO
; APPLICANT: SHINOZAKI, KAZUO
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
; TITLE OF INVENTION: CLEAVAGE ENZYME GENE
; FILE REFERENCE: 3914-3
; CURRENT APPLICATION NUMBER: US/09/758,269
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: JP 2001-003476
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: JP 2000-010056
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: Vigna unguiculata
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1836)
US-09-758-269-11
Alignment Scores:
Pred. No.:      3,536-239      Length:      1839
Score:          2168.50      Matches:      423
Percent Similarity: 78.44%      Conservative: 61
Best Local Similarity: 68.56%      Mismatches: 98
Query Match:      68.84%      Indels:      35
DB:              10          Gaps:        8
US-09-758-269-6 (1-599) x US-09-758-269-11 (1-1839)
Qy      8 AlaAlaValSerGlyArgTyrLeuGlyGlyAsnHisThrGlnProLeuSerSerSer 27
||||:|||||
Db      7 TCATCAGCTTCAACACTTGGTTTAACGCCACACTCCCATCTCCCCCTTCAAGACCTA 66
||||:|||||
Qy      28 GlnSerSerAspLeuSerTyrCysSerSerLeuProMetAlaSerArgValThrArgLys 47
||||:|||||
Db      67 CCTTCCACATCT--TCTCCACAAACTTACTTCTTTA-----AGGAAACATCCCTCT 117
||||:|||||
Qy      48 LeuAsn---ValSerSerAlaLeuHisThrProProAlaLeuHisPheProLysGln--- 65
||||:|||||
Db      118 TCCAACACCATCACATGTTCCCTTCAACA-----CTCCACTTCCCAACACAGTAC 168
||||:|||||
Qy      66 -----SerSerAsnSerProAlaValValLysProLysAlaLysGluSer 81
||||:|||||
Db      169 CAACCAACATCCACATCCACATCCACAGCCACCACCAACACCCCACTCAAAACT 228
||||:|||||
Qy      82 AsnThr----- 83
||||:|||||
Db      229 ACCACCATCACACCACACCGCCAGGGAAACCAACCCCTCTCTCTGACACCAACCAA 288
||||:|||||
Qy      84 -----LysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAla 100
||||:|||||
Db      289 CCATTACCTCAAAATGGAACCTTTCTCCAGAAAGCGCTGCCACGGCCCTTGACCTGGTC 348
||||:|||||
Qy      101 GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer 120
||||:|||||
Db      349 GAAACGGCGCTGCTTCGCACGAGCGCAACACCCCGCTCCCAAAACCGCGGACCCGAGG 408
||||:|||||
Qy      121 ValGlnIleAlaGlyAsnPheAlaProValAsnGlnProValArgArgAsnLeuPro 140
||||:|||||
Db      409 GTCCAAATCGCGGGAACTTCGCGCGCGTGGCGGAGCATGCGCGGAGCATCAAGACTCCCG 468
||||:|||||
```

```
Db 1546 TGGCCCAAGTCTCGGCGCTTGGCAAGAGTTGATTTCTGAGTGGGGAAGTGAAGAAGTAC 1605
Qy LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGlyGlu 540
Db 1606 ATGTATGGAGAGAGAGTTCGCTGGGAGCCTCTGTTCTTCCC---AACGGCCAAA 1662
Qy GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560
Db 1663 GAAGACGATGGGTATATTCTGGCATTCGTGCACGACGAGAAAGAAATGGAATCCGAGCTG 1722
Qy 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580
Db 1723 CAGATTGGAATGCCCAAAATTTAAAGCTCGAAGCTTCATCAAACTCCCTCTCGTGT 1782
Qy 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGln 597
Db 1783 CCTACGGTTTTCATCGAAGTTTCATTCATCCAGGATTTGAGGAACAA 1833

RESULT 4
US-09-758-269-1
; Sequence 1, Application US/09758269
; Patent No. US20020104120A1
; GENERAL INFORMATION:
; APPLICANT: IUCHI, SATOSHI
; APPLICANT: KOBAYASHI, MASATOMO
; APPLICANT: SHINOZAKI, KAZUO
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
; TITLE OF INVENTION: CLEAVAGE ENZYME GENE
; FILE REFERENCE: 3914-3
; CURRENT APPLICATION NUMBER: US/09/758,269
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: JP 2001-003476
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: JP 2000-010056
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1749)
US-09-758-269-1

Alignment Scores:
Pred. No.: 9.18e-219 Length: 1752
Score: 1991.00 Matches: 380
Percent Similarity: 78.44% Conservative: 82
Best Local Similarity: 64.52% Mismatches: 105
Query Match: 63.21% Indels: 22
DB: 10 Gaps: 6

US-09-758-269-6 (1-599) x US-09-758-269-1 (1-1752)
Qy 16 GlyGlyAsnHisThrGlnProLeuSerSerGlnSerSerAspLeuSerTyr--- 34
Db 31 GGTGGTATTAAACATGGCCT-----CAGGCCCAATTTGATTTGGGTTTATAG 78
Qy 35 ---CysSerSerLeuProMetAlaSerArgValThrArgLysLeuAsnValSerAla 53
Db 79 CCATTAAAGAACCAACGAAGTTATTAAATGCACGGTGCAGATCGACGTAACGAA--- 135
Qy 54 LeuHisThrProProAlaLeuHisPheProLysGlnSerSerAsnSerProAlaIleVal 73
Db 136 TTAACCAAAAAACGCCAATTTATTACCCAGAACCCCGCTACTCCGCGC----- 186
Qy 74 ValLysProLysAlaLysGluSerAsnThrLysGlnMetAsnLysPheGlnArgAla 93
Db 187 -----CAGCATATCTCTCCGCTAAACATCTTCAGAAAGCGCG 228
```

```
Qy 94 AlaAlaLeuAspAlaGluGlyPheLeuValSerHisGlyLysLeuHisProLeu 113
Db 229 GCGATTGGATCGACGCGCTGAGCGTGCATTAATCTCACAGCAAGATTCTCCACTT 288
Qy 114 ProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValAsnGln 133
Db 289 CCAAAACCGCTGATCCAGCTGTTCCAGATTGCGCGGAATTATTCGCCCGTACCGGAATCT 348
Qy 134 ProValArgAsnLeuProValValGlyLysLeuProAspSerIleLysGlyValTyr 153
Db 349 TCCGTCGCGGAAACCTCACCGTCGAAGAAACATCCCTGACTGCTGATTCACCGGTGTAT 408
Qy 154 ValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHisPhePheAspGly 173
Db 409 ATCCGTAACGCGCGCAATCCGATGTTGAGCCACAGCTGGGCACCATTTATTCGACGGA 468
Qy 174 AspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyrAlaCysArgPhe 193
Db 469 GACGGAATGGTTTCACGCAAGTTAAACCAACCGGTTACGCTAGCTACGCATCGCGGTTT 528
Qy 194 ThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPheProLysAla 213
Db 529 ACAAAACCGAGAGATTGGTTCAGAAACACGATTTGGTTCGACCAAGTTTCCCGAAGCA 588
Qy 214 IleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAla 233
Db 589 ATCGCGGAGCTTCACGGTCACTCGGGAATCGCAGCTTTGATCTGCTTTTACGCACGTGG 648
Qy 234 AlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyr 253
Db 649 CTTTGTGTCTGATCAACCAACCGCGCTCGGAGTAGCAACCGCGTTTGGTTTAC 708
Qy 254 PheAsnGlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIleThr 273
Db 709 TTTAATAACCGCTTTTAGCTATGTACAGACGATTTACCGTACCAATTTAAATTAATCT 768
Qy 274 ProAsnGlyAspLeuLysThrValGlyArgPheAspPheAspGlyGlnLeuGluSerThr 293
Db 769 CAACCGCGCATCTCCAAACCGTTGGACGTTACGATTTTCGACGGTCAGTTAAATCCGCA 828
Qy 294 MetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeuSerTyrAsp 313
Db 829 ATGATAGCTACCCGAAACTGGACCGCTTACGAAGAGCTTTCACGCGTTAAGTACGAC 888
Qy 314 ValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSerPro 333
Db 889 GTCGTTAAGAAACCTTACTCTGAATACTTTCAGATTCTCCACAGCGGCTTAAATCCGCG 948
Qy 334 AspValGluIleGlnLeuAspGlnProThrMetHisAspPheAlaIleThrGluAsn 353
Db 949 GAAATGGAGATCCCGCTCGAAACTCCGACGATGATTACGATTTTCGCTATTAACGAGAAT 1008
Qy 354 PheValValValProAspGlnValValPheLysLeuProGluMetIleArgGlyGly 373
Db 1009 TTTGTGGTGAATCTCTGATCAACCAAGTCGTGTTTCAAGCTCGCGAGATGATTTCCGCTAA 1068
Qy 374 SerProValValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAla 393
Db 1069 TCTCCGGTGTGTTTCGACGGAAGAAAGGTTTCCCGATTTGGGATTAATCCCAAGGACGG 1128
Qy 394 GluAspSerSerAsnIleLysTrpIleAspAlaProAspCysPheCysPheHisLeuTyr 413
Db 1129 ACAGAAGCTTCTCAGATAATCTGGGTGAACCTCCGGAGACGCTTCTGTTTTCATCTCTGG 1188
Qy 414 AsnAlaTrpGluProGluThrAspGluValValIleGlySerCysMetThrPro 433
Db 1189 AATGCATGGGAATCGCGGACGAGAGAGATTTGTGTGATCGGATCGTATGTGCGCG 1248
Qy 434 ProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArg 453
Db 1249 CGCGATTCAATCTTCAACGAGAGACGAGAGCTTTGAGAAGCGTTTTCGAGATCAGG 1308
Qy 454 LeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnGln 473
```

```
Db 1309 ATAAACCTCAGACACGTAACACACGCGTTCGTTGGTTGGTTAAACGAGAT----- 1362
Qy 474 ValAenLeuGluAlaGlyMetValAenArgAenMetLeuGlyArgLysThrLysPheAla 493
Db 1363 GTAAATTTAGAGATTGGTATGTTAAACCGGAACCGGTTAGGAAGAAAAACCCGGTTCCGC 1422
Qy 494 TyrLeuAlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAspLeuThr 513
Db 1423 TTTTGGCTATGCTTATCTTGGCCAAAGTTTCCCGGTTTCGGCTTAAGTTCGATCTTTC 1482
Qy 514 ThrGlyGluValLysLysHisLysThrGlyAspAenArgTyrGlyGlyGluProLeuPhe 533
Db 1483 ACCGGTAGATGAAAAAATATATTACCGCGTGAGAAATATGCGGCGAACCCTTTTTC 1542
Qy 534 LeuProGlyGluGlyGly-----GluGluAspGluGlyTyrIleLeuCysPhe 549
Db 1543 TTGCCCGCAACTCCGGTAACGGCGGAAGAAATGAAGATGACGGTTATATATTGTCAC 1602
Qy 550 ValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAenAlaValSerLeuGlu 569
Db 1603 GTTCATGACGAAGAAACAAAGACATCAGAGCTTCAGATTATTAACGCTGTTAATTAAAG 1662
Qy 570 ValGluAlaThrValLysLeuProSerArgValProTyrGlyPheHisGlyThrPheIle 589
Db 1663 CTTGAGCTACGATTAAACTACCGCTCTAGATACCGTATGCGTTTCATGGCACATTGTCG 1722
Qy 590 GlyAlaAspLeuAlaLysGlnVal 598
Db 1723 GATTGGAATGAACCTCGTTGATCAATTA 1749
```

## RESULT 5

```
US-09-758-269-13
; Sequence 13, Application US/09758269
; Patent No. US20020104120A1
; GENERAL INFORMATION:
; APPLICANT: IUCHI, SATOSHI
; APPLICANT: KOBAYASHI, MASATOMO
; APPLICANT: SHINOZAKI, KAZUO
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
; FILE REFERENCE: CLEAVAGE ENZYME GENE
; CURRENT APPLICATION NUMBER: US/09/758,269
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: JP 2001-003476
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: JP 2000-010056
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1815
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1812)
US-09-758-269-13
```

## Alignment Scores:

Pred. No.:	1,03e-211	Length:	1815
Score:	1930.00	Matches:	375
Percent Similarity:	75.68%	Conservative:	73
Best Local Similarity:	63.34%	Mismatches:	122
Query Match:	61.27%	Indels:	22
DB:	10	Gaps:	8

US-09-758-269-6 (1-599) x US-09-758-269-13 (1-1815)

Qy 22	ProProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMetAla 41
Db 49	CCGGCCCGGTCAGGCGCCGGGCGCTCC-----AATTCCGTCAGGTTCTCG 93

```
Qy 42 SerArgValThrArgLysLeuAenValSerSerAlaLeuHisThrPro-----Pro 58
Db 94 CCGCGCCCGTCAGCTCCGTCGCGCCCGCGAGTGCTCTCAGGGCGCGTTCACCAAGCCC 153
Qy 59 AlaLeuHisPheProLysGlnSerSerAsnSerProAlaIleValVal-----74
Db 154 GTCCGCCACCTGCCTCGCGCTCCAGGAAGCCCGCCCATTTGCCGTCACGAGGCACGCC 213
Qy 75 ---LysProLysAlaLysGluSerAsnThrLysGlnMetAsnLeuPheGlnArgAlaAla 93
Db 214 GCGCGCCGCGAAGCGGAGCGGCGCAAGAAGCAGCTCAACTTGTTCAGCGCGCGCGG 273
Qy 94 AlaAlaAlaLeuAspAla---AlaGluGlyPheLeuValSer---HisGluLysLeuHis 111
Db 274 CCGCGCCCGCTCGACGCGTTTCGAGGAAGGGTTCTGTGCCAACGTCCTCGAGGCGCCAC 333
Qy 112 ProLeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValAsn 131
Db 334 GGGCTGCCAGCAGCGCGCGCGCGTGCAGATCGCGCAACTTCGCGCCCGTCGGG 393
Qy 132 GluGlnProValArgAsnLeuProValValGlyLysLeuProAspSerIleLysGly 151
Db 394 GAGAGCGCGCGCGTGCAGCAGCTCCCGCTCTCCGCGCGCATCCCGCCCTTCATCGAGCGG 453
Qy 152 ValTyrValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHisHisPhePhe 171
Db 454 GTCTACCGCGCAACGGCGCCACCCCTCTGACCCCGCTCGCGGGGCGCACCTCTTTC 513
Qy 172 AspGlyAspGlyMetValHisAlaValLysPheGluHisGlySerAla---SerTyrAla 190
Db 514 GACGGCAGCGCATGTTGTCAGCGCTCGCGATACGCAACGCGCGCGCGAGTCTCTACGCC 573
Qy 191 CysArgPheThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPhe 210
Db 574 TGCCGCTTCACGAGACCGCGCGCTCGCGCAGAGCGCGCATCGCGCGCGCGCTCTTC 633
Qy 211 ProLysAlaIleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyr 230
Db 634 CCCAAGGCCATTGGCGAGCTGCACGGGACCTCCGGGATCGCGCGCTTCGCCCTGTTCTAC 693
Qy 231 AlaArgAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAlaAsnAlaGly 250
Db 694 GCGCGCGCGCGTGCAGCGCTCTGTGACCCCTCGCGCGCGCACCGCGCGTGGCCACCGCGG 753
Qy 251 LeuValTyrPheAsnGlyArgLeuAlaMetSerGluAspAspLeuProTyrGlnVal 270
Db 754 CTCGCTCTACTTCAACGGCGCGCTGCTCGCCATGTCGAGGACGACCTCCCTACACCGTTC 813
Qy 271 GlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAspPheAspGlyGlnLeu 290
Db 814 CCGTGGGCGGACGCGGCGACCTCGAGACCGTCGCGCGCTACGACTTCGACGGGCGAGCTC 873
Qy 291 GluSerThrMetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeu 310
Db 874 GGCTGGCCCATGATCGCGCACCCCAAGCTGAGCCCGCGCACCGGGGAGTCCACCGCGCTC 933
Qy 311 SerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThr 330
Db 934 AGCTACGACGTCATCAAGAGCGCGTACTCTCAAGTACTTCTACTTCAGCGCCCGCGGACCC 993
Qy 331 LysSerProAspValGluIleGlnLeuAspGlnProThrMetHisAspPheAlaIle 350
Db 994 AAGTCCGACGAGTGGAGATCCCGCTGGAGCAGCCCGACGATGATCCAGCACTTCGCCATC 1053
Qy 351 ThrGluAsnPheValValProAspGlnGlnValValPheLysLeuProGluMetIle 370
Db 1054 ACCGAGAACTTCGTGGTTGTGCCCGCACCCAGGTGGTTCAGCTTCAAGCTCCAGGAGATGCTG 1113
Qy 371 ArgGlyGlySerProValValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAsp 390
Db 1114 CCGCGCGCGTCCCGCGGTGGTCTGCAAGAGAGACGTCGCGGTTTCGGGCTGCTCCCC 1173
Qy 391 LysTyrAlaGluAspSerSerAsnIleLysTrpIleAspAlaProAspCysPheCysPhe 410
```

```

1174 AACGACGCCGCGGACGCGTGGAGATGGGTGGAGCGTCCGACGCTCTCTCTTC 1233
411 HisLeuTrpAsnAlaTrpGluGluProGluThrAspGluValValValleGlySerCys 430
1234 CACCTGTGGAACGCGTGGAGGACGAGCGGCGGAGGTGGTGTGTGCTCTCTGC 1293
431 MetThrProProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSer 450
1294 ATGACCCCGCGCGACTCCATCTTCAACGAGTCCGAGCGCGCTGGAGCGGTGTGAC 1353
451 GluIleArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGlu 470
1354 GAGATCCGCTGGACCGCGCGACGCGCGTCCACGCGCGCGCTCTCTCTCTCTCTCG 1410
471 AspGlnGlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThr 490
1411 TCGCAGCAGGAGAACCTGGAGGTGGCATGGTGAACCGCAACCTGCTGGCGCGGAGAC 1470
491 LysPheAlaTyrIleuAlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysVal 510
1471 CGGTAGCGGTACTCTCGCGTGGCGGAGCGGTGGCCCAAGGAGTTCGGGCTTCGCAAG 1530
511 AspLeuThrThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGlu 530
1531 GACCTGTCCAGCGCGGAGCTCACCAGTTCGAGTACGGCGGCGCGGTTCGGCGCGGAG 1590
531 ProLeuPheLeuProGlyGlyGlyGlu-----GluAspGluGlyTyr 545
1591 CCCTGCTTCTCCATGACCCGCGCGCGGCCGCCCGCGCGGCGGAGACGCGGTAC 1650
546 IleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAla 565
1651 GTGCTCACCTTCGTCACGACGAGCGCGCGCGACGTCGAGTACTTGTGTCAATGCC 1710
566 ValSerLeuGluValGluAlaThrValLysLeuProSerArgValProTyrGlyPheHis 585
1711 GCCGACATCCGGTGGAGCCACGGTTCAGTCCGCGTCCCGCTCCCGCTTCGGCTTCCAC 1770
586 GlyThrPheIleGlyAlaAspLeuAlaLysGln 597
1771 GCCACCTTCATCAGCGCGGCGGAGCTCGAGGCCAG 1806

```

## RESULT 6

```

US-09-758-269-9
; Sequence 9, Application US/09758269
; Patent No. US20020104120A1
; GENERAL INFORMATION:
; APPLICANT: IUCHI, SATOSHI
; APPLICANT: KOBAYASHI, MASATOMO
; APPLICANT: SHINOZAKI, KAZUO
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
; TITLE OF INVENTION: CLEAVAGE ENZYME GENE
; FILE REFERENCE: 3914-3
; CURRENT APPLICATION NUMBER: US/09/758,269
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: JP 2001-003476
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: JP 2000-010056
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1731)
US-09-758-269-9

```

Alignment Scores:

Pred. No.: 4.83e-181 Length: 1734

```

Score: 1663.50 Matches: 323
Percent Similarity: 72.56% Conservative: 100
Best Local Similarity: 55.40% Mismatches: 137
Query Match: 52.81% Indels: 23
DB: 10 Gaps: 9

US-09-758-269-6 (1-599) x US-09-758-269-9 (1-1734)

QY 24 LeuSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMetAlaSerArg 43
Db 25 CTTCTCCGACGACGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 84
QY 44 ValThrArgLysLeuAsnValSerSer---AlaLeuHisThrProProAlaLeuHisPhe 62
Db 85 ATTCTCTGACGAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 144
QY 63 ProLysGlnSerSerAsnSerProAlaIleValValLysProLysAlaLysGluSerAsn 82
Db 145 CCGGTTCGCTCACCG-----GTTAAGCTCAAACCAACGATATCCAAAC 186
QY 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaAlaLeuAspAlaAlaGlu--- 101
Db 187 -----TTAAACCTTCTTCAGAGCTAGCGGTCTAGATCTCGACAGATTGAGTCC 237
QY 102 GlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerVal 121
Db 238 TCTATCGTTATTCCTATGAGCAGAAATCGCGCTTCTCTCTCTCTCTCTCTCTCTCT 297
QY 122 GlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgAsnLeuProVal 141
Db 298 CAATTATCAGGTAACTTCGCTCCGTTAATGAATCTCCGGTTCAGAACGGTTTAGAAGTG 357
QY 142 ValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyValAlaAsnProLeu 161
Db 358 GTTGTCTAGATTCCTTCTGTCTAAAGAGGTTTACATCCGTAAACCGTCAACCCCTATG 417
QY 162 HisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaValLys 181
Db 418 TTTCCGCGTGTAGCCGACATCATTTATTTGACGCTGACGGAATGATTTCACCGCTTAGT 477
QY 182 -----PheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPhe 199
Db 478 ATCGGTTTGTATAAC---CAGGTTAGTTACAGCTCCGCTACACTAAACCAACCGCGTT 534
QY 200 ValGlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGly 219
Db 535 GTTCAAGAAACCGCGCTTGGACGATCGGTTTCTCTAAACCAACATCCGCGAGGTTCCAGGC 594
QY 220 HisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAsp 239
Db 595 CATTCCGCTCTAGCTCGACTCGCTCTCTCTACGCGCTCGAGCTGGGATCGGTCTAGTGGAC 654
QY 240 ProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeu 259
Db 655 GGGACACGTCATGGCGGTAGTAAACCGCGGTGTGGTTTCTTTTAAACCGGAGGTTATTAT 714
QY 260 AlaMetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLys 279
Db 715 GCCATGTCAGAGATGATCTCTCTTACCAAGTGAAGATCGAGCTGCAAGGAGATCTTGAG 774
QY 280 ThrValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLys 299
Db 775 ACGATCCGACGCTCGGATTCGATCCAGCAGATTTGACTCTTCAGTGATAGCGCATCTTAAG 834
QY 300 ValAspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyr 319
Db 835 GTGGACGCGACACGAGGAGATCTCCATACACTGAGCTACAACGTTTGAAGAAACCTCAT 894
QY 320 LeuLysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeu 339
Db 895 CTCAGGTATCTTAATTCACACGTCGCGGAAAAGACACGTCGCTGGAGATCACGCTC 954
QY 340 AspGlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAsp 359

```

```
Db 955 CCTGAACACGATGATTCATGATTTCCGATACCGAGAAATTTGTCGTTATACCGAT 1014
Qy 360 GlnGlnValValPheLysLeuProGluMetIleArgGlySerProValValTyrAsp 379
Db 1015 CAGCAATGGTATTCAAAATATCCAAATGATTCGGGGCGGTCACCGGTTACTACGTT 1074
Qy 380 LysAsnLysValAlaAtpPheGlyIleLeuAspLysTyrAlaGluAAspSerSerAsnIle 399
Db 1075 AAAGAAAATATGGCGAGATTTGGAGTTTGTCAAGCAGGATCTGACCGGTCGGATATA 1134
Qy 400 LysTrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluPro 419
Db 1135 AATTGGGTGATGATACCGGATGTTCTTCGTTCCATCTATGGAATGCGTGGGAAGAG-- 1191
Qy 420 GluThrAspGlu-----ValValIleGlySerCysMetThrProProAsp 435
Db 1192 AGAACCGAAGAGGAGACCCAGTATATCGTCTGTAATCGGGTCATGTATGAGCCACCCGAC 1251
Qy 436 SerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsn 455
Db 1252 ACGATCTTTAGTGAATCAGGAGAACCAACCCGGGTTGAAATTAAGTGAGATCCGGTTAAAC 1311
Qy 456 LeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsn 475
Db 1312 ATGGGTACAAAGAAATCGAACGTAAGTTATCGTAACCGGA-----GTGNAAT 1359
Qy 476 LeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeu 495
Db 1360 TTAGAAGCGGGTCACATAAACCGTAGTTACGTGGCGCGGAAAGCCAGTTCGTTTACATA 1419
Qy 496 AlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAspLeuThrThrGly 515
Db 1420 GCAATAGCCGATCTTGGCCCAAAATGCGAGTGCGCATTCGGAAGGTAGATATACAAACCGC 1479
Qy 516 GluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGluProLeuPheLeuPro 535
Db 1480 ACCGTTTCAGAGTTTAATACGGACCGAGCGGTCGTTGGTGGAACCGTCTTGTACCG 1539
Qy 536 GlyGluGlyGlyGluGluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThr 555
Db 1540 GAGGAGAGAGGAGAGAGAGAAAGAGTTATGTAATGGGGTTTGTGAGAGAGCAAGAGAAA 1599
Qy 556 TrpLysSerGluLeuGlnIleValAsnAlaValSerLeuGluValGluAlaThrValLys 575
Db 1600 GACGAGTCGGAGTTTGTGCGTGCACGCGACGGATATGAAGCAAGTCGCGCGGTGCGC 1659
Qy 576 LeuProSerArgValProTyrGlyPheHisGlyThrPheIleGlyAlaAspLeuAla 595
Db 1660 TTGCGGAGAGGGTACCTTATGGTTTCCATGGAGACGTTTCGTGAGCGAGAAATCAGTTGAAG 1719
Qy 596 LysGlnVal 598
Db 1720 GAACAAGTT 1728
```

## RESULT 7

```
US-09-758-269-3
; Sequence 3, Application US/09758269
; Patent No. US20020104120A1
; GENERAL INFORMATION:
; APPLICANT: TUCHI, SATOSHI
; APPLICANT: KOBAYASHI, MASATOMO
; APPLICANT: SHINOZAKI, KAZUO
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
; TITLE OF INVENTION: CLEAVAGE ENZYME GENE
; FILE REFERENCE: 3914-3
; CURRENT FILING DATE: 2001-01-12
; PRIOR FILING DATE: 2001-01-12
; PRIOR FILING DATE: 2001-01-11
; PRIOR FILING DATE: 2001-01-11
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1785)
US-09-758-269-3
Alignment Scores:
Pred. No.: 5,15e-101 Length: 1788
Score: 969.00 Matches: 214
Percent Similarity: 55.31% Conservative: 114
Best Local Similarity: 36.09% Mismatches: 221
Query Match: 30.76% Indels: 44
DB: 10 Gaps: 13
US-09-758-269-6 (1-599) x US-09-758-269-3 (1-1788)
Qy 32 LeuSerTyrCysSerSerLeuProMetAlaSerArgValThrArgLysLeuAsnValSer 51
Db 56 CTCTTCTTCCGCCGATCTTCTCTC-----CTACTC 88
Qy 52 SerAlaLeuHisThrProAlaLeuHisPheProLysGlnSer-----SerAsn 68
Db 89 TCCTCCGTATCACTCCGCGTCTCGAAGACGTTCTCCATCACAACCCAGCGACA 148
Qy 69 SerProAlaIleValLysProLys-AlaLysGluSerAsnThrLysGlnMetAsnLe 88
Db 149 ACAATGATCGTCGTAACAACCCCAACACCTCCACAAACCAACCAATCACAACCTTAGTCT 208
Qy 88 uPheGlnArgAlaAlaAlaAlaLeuAspAla-AlaGluGlyPheLeuValSerHisG 108
Db 209 CATCACCCAGAACTCCGACCAAGAAATGACTCTCGCAACAGCTCTTTCACCCCGTCG 268
Qy 108 LuLysLeu-----HisProLeuProLysThrAlaAspProSerValG 122
Db 269 AAGATGTAATCAACACGTTTCATCGATCCACTTCACGTCCTTCCTGTTGATCCAAAACATG 328
Qy 122 lnIleAlaGlyAsnPheAlaProValAsnGlnProValArgArgAsnLeuProValV 142
Db 329 TCCTCTCTGATAACTCTCGCTCTCGACGAGCTTCTCTCCACACAGACTGTGAAATCA 388
Qy 142 al---GlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnProL 161
Db 389 TCCACGCGACTCTTCCACTGTCACTTAACGCGGCTTACATCCGTAACGGTCCAAATCCAC 448
Qy 161 euHisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaValL 181
Db 449 AGTTTCTCCCTCGTGGTCCTTACCATCTCTTCGACGGCGGAGCGGTATGCTTCACGCCATAA 508
Qy 181 ysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheValG 201
Db 509 AAATCCCAACACGGTAAAGCCACTCTCTGTAGCAGATACGTCAAGACTTATAAATACAACG 568
Qy 201 lnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHisT 221
Db 569 TCGAGAAACAAACCGGAGCTCCGGTTATGCCCTAACTGTTTCCCGGATTCACCGGTGTA 628
Qy 221 hr---GlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspP 240
Db 629 CGGCGTCAGTAGCTCGTGGAGCTTTAAGCGGAGCTAGGTTTAAACCGGACAGTATAATC 688
Qy 240 roAlaHisGlyThrGlyValAlaAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuA 260
Db 689 CGGTTAACGGCATTTAGCTAATAACAAGCTAGCTTTCTTCAGTAACCGTCTCTTTG 748
Qy 260 laMetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysT 280
Db 749 CTTTAGGTGATCTGATTTACCTTACCGTCCGCTACCGGATTAACCGAATCAGGAGATTTGAAA 808
Qy 280 hrValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysV 300
```



```
Db 809 CGATCGACGGTACGATTTTCGACGGAAATTCAGATGAGTATGACAGCTCATCCCTAAAA 868
Qy 300 aAspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrL 320
Db 869 CCGATCCAATAACCGAGAAACTTCGCTTCGGTACGGTCCGGTT---CCACCGTTT 925
Qy 320 euLysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluLeu---GlnL 339
Db 926 TACATATTTTCGGTTTGGATTCGCGCGGAAACAAAGAGACGCTTCGATATTCGCA 985
Qy 339 euAspGlnProThrMetMetHisAspPheAlaLeuThrGluAsnPheValValProA 359
Db 986 TGAGTCTCGTTCGTTTCCATGACTTCGGATCACAAACGCTACGGGATTTTCGCGAG 1045
Qy 359 spGlnGlnValValPheLys-----LeuProGluMetIleArgGlyGlySerProV 376
Db 1046 AGATTCAGCTTGCATGAGGATGAACATGTTGGATTTGTTCTCGAAGGTGGTTCTCCG 1105
Qy 376 alValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAsp 396
Db 1106 TTGGTACTGATACGCGAAACTCCAGGCTTCGAGTGATTCCTAAGTACGCGCGGAGATG 1165
Qy 396 erSerAsnIleLysTrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAla 416
Db 1166 AGTCGAGATGAATGTTTCGAAGTTCCTGGATTCATATCATTCACGCTATTAATGCTT 1225
Qy 416 rpGluGluProGluThrAspGluValValIleGlySerCysMetThrProProAsp 436
Db 1226 GGGATGAAGATGATGAACACAGCGCTGTTTGTATTCACCGCAATATATGTCGATGAAC 1285
Qy 436 erIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnL 456
Db 1286 ATACTTTAGAGAGATGGAT---CTGGTTCATGCTTTGGTGGAGAGGTGAAGATCGATC 1342
Qy 456 euLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnL 476
Db 1343 TCGTACCAGGATGTGAGACGTCATCCGATCTCAGCGAGG-----AATC 1387
Qy 476 euGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuA 496
Db 1388 TCGATTTCCGCTGTGATTATCCGCGCTTCTCGGAGATGTAGCAGGTACGTTTACGCG 1447
Qy 496 laLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyG 516
Db 1448 CGATTGGAGATCCGATGCGCAAGATCTCGGTGTGGTGAAGCTTGATGTCTAAAGGAG 1507
Qy 516 lu-----ValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGluP 531
Db 1508 ATCGGATGATTGATCGGTGGCCCGGAGAAATGTACGGTTTCAGGTTGTTACGCGCGGAG 1567
Qy 531 roLeuPheLeuProGlyGluGlyGly-----GluGluAspGluGlyTyrIleL 547
Db 1568 CGTTTTCGTAGTAGGATCTGGTAATCCGGAGCGCGAGAGAGATGATGTTATGTGG 1627
Qy 547 euCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaVal 567
Db 1628 TCACGTATGTTCCAGATGAAGTACTGGAGAAATCGAAGTTTCTGGTGATGGACGCTAAAT 1687
Qy 567 er-----LeuGluValGluAlaThrValLysLeuProSerArgValProTyrGlyPheH 585
Db 1688 CCGCGAGCTTGAATTCGTCGCGCGGTGAGGTTCGCGGAAGGGTTCGTACGAGATTCC 1747
Qy 585 isGlyThrPheIleGlyAlaAspAspLeuAlaLys 596
Db 1748 ATGGGTATTGTCAGGAAAGTGACCTTAATAAG 1782
```

## RESULT 8

```
US-09-938-842A-1444
; Sequence 1444, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
```

```
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1444
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-938-842A-1444
```

## Alignment Scores:

Pred. No.:	5,15e-101	Length:	1788
Score:	969.90	Matches:	214
Percent Similarity:	55.31%	Conservative:	114
Best Local Similarity:	36.09%	Mismatches:	221
Query Match:	30.76%	Indels:	44
DB:	10	Gaps:	13

US-09-758-269-6 (1-599) x US-09-938-842A-1444 (1-1788)

```
Qy 32 LeuSerTyrCysSerSerLeuProMetAlaSerArgValThrArgLysLeuAsnValSer 51
Db 56 CTCTTCTTCGCGCGGATCTTCTCTCTC-----CTACTC 88
Qy 52 SerAlaLeuHisThrProProAlaLeuHisPheProLysGlnSer-----SerAsn 68
Db 89 TCTCTCGTATCAACTCCGCGCTGCGAAGAACGTTCTTCTCAATCAAAACCAAGCGACA 148
Qy 69 SerProAlaIleValValLysProLys-AlaLysGluSerAsnThrLysGlnMetAsnLe 88
Db 149 ACAATGATCGTGTGAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 208
Qy 88 uPheGlnArgAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 108
Db 209 CATCACCCGAAATCTCCGACGAGAAATGACTCTCGCAACAGCTCTCTTCCACCGCTG 268
Qy 108 luLysLeu-----HisProLeuProLysThrAlaAspProSerValG 122
Db 269 AAGATGTAATCAACACAGTTCATCGATCCACCTTTCACGCTTTCGCTTGATCCAAACATG 328
Qy 122 InIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuProVal 142
Db 329 TCTCTCTGTAACTTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 388
Qy 142 al---GlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnProL 161
Db 389 TCCACGGCACTCTTCCACTGTCTTAAACGGCGCTTACATCGTAAACGGTCCAAATCCAC 448
Qy 161 euHisGluProValThrGlyHisHisPheAspGlyAspGlyMetValHisAlaValL 181
Db 449 AGTTTCTCCCTCGTGGTCTTACCATCTCTTCGACGGGACGATGCTTCCACGCGATAA 508
Qy 181 ysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheValG 201
Db 509 AAATCCCAACGGTAAGCCACTCTCTGTAGCAGATACGTCAAGACTTATAAATCAACG 568
Qy 201 InGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 221
Db 569 TCGAGAAACAAACCGGAGCTCCGGTTATGCTTAACGCTTTTCCGGATTCACCGGTGTA 628
Qy 221 hr---GlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaAlaGlyIleValAsp 240
```

```
Db 629 CGGCGTCAGTAGCTCGTGAGGCTTTAACGGCAGCTAGGTTTAAACCGGACAGATATAATC 688
Qy 240 roAlaHisGlyThrGlyValAlaAenAlaGlyLeuValTyrPheAsnGlyArgLeuLeuA 260
Db 689 CGGTTAACGGCATTCGGTTAGCTATAACAAGCTAGCTTCTTCAGTAACCGCTCTCTTTG 748
Qy 260 laMetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysT 280
Db 749 CTTTAGGTGAATCTGATTTACCTACGGCGTCGGAATTAACGAATCAGAGATATTGAAA 808
Qy 280 hrValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysV 300
Db 809 CGATCGGACGGTACGATTTTCGACGGAAATAGCGATGAGTATGACAGCTCATCTTAAA 868
Qy 300 alAspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrL 320
Db 869 CCGATCCCAATAACCGGAGAAATTTTCGCTTTTCGGTACGGTCCGGT---CCACCGTTT 925
Qy 320 euLysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIle---GlnL 339
Db 926 TAACATATTTCCGGTTTGAATTCGCGCGGGAATAAAACAAAGAGACGTTCCGATATTCG 985
Qy 339 euAspGlnProThrMetHisAspPheAlaIleThrGluAsnPheValValProA 359
Db 986 TGACGCTCCGTCGTTTCTCCATGACTTCGCGATCACGAACGTCACGCGATTTTCGCAG 1045
Qy 359 spGlnGlnValValPheLys-----LeuProGluMetIleArgGlyGlySerProV 376
Db 1046 AGATTTCAGCTGGCATGAGGATGAACATGTTGGATTTGGTTCTCGAAGTGGTTCTCCGG 1105
Qy 376 alValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAsp 396
Db 1106 TTGGTACTGATAACCGGAATAAACTCAAGCGTTGGAGTGAATTCCTAAGTACCGCGAGATG 1165
Qy 396 erSerAsnIleLysTrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaT 416
Db 1166 AGTCGGAGATGAATGGTTTCGAAGTTCCTGGATTCATATCATTCACCGCTTATATGCTT 1225
Qy 416 rpGluGluProGluThrAspGluValValValIleGlySerCysMetThrProProAsps 436
Db 1226 GGGATGAAGATGATGAAACACGCGCTGTTTGAATTCACCGAATATATGTCGATTTGAAC 1285
Qy 436 erIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAenL 456
Db 1286 ATACTTTAGAGAGATGATG---CTGGTTCATGCTTTCTGGAGAGGTGAAGATCGATC 1342
Qy 456 euLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnL 476
Db 1343 TCGTCACCGGGATTGTGAGACGTCATCCGATCTCAGCGAGG-----AATC 1387
Qy 476 euGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuA 496
Db 1388 TCGATTTTCGCTGTGATTAATCCGGCGCTTTCGGGAGATGTAGCAGGTACGTTTACGCGG 1447
Qy 496 laLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyG 516
Db 1448 CGATTGGAGATCCGATGCGGAAGATCTCCGGTGTGGAAGCTTGATGTGCTAAAGGAG 1507
Qy 516 lu-----VallysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGluP 531
Db 1508 ATCCGGATGATTGTACGTTGGCCCGCTAGAATGTACGGTTTCAGGTTGTACGCGGAGAAC 1567
Qy 531 roLeuPheLeuProGlyGluGlyGly-----GluGluAspGluGlyTyrIleL 547
Db 1568 CGTTTTCGTAGCTAGGATCTCGTGAATCCGGAGCGGAGGAGATGATGTTATGTGG 1627
Qy 547 euCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaValS 567
Db 1628 TGACCATATGTTCCAGCATGAAGTACTGGAGAAATCGAAGTTTCTGGTGAAGCAGCTAAAT 1687
Qy 567 er-----LeuGluValGluAlaThrValLysLeuProSerArgValProTyrGlyPheH 585
Db 1688 CCGCGGAGCTTGAAATCGTCGCGCGCGTGAGGTTCCCGGGAAGGTTCCGTACGAGTTCC 1747
```

```
Qy 585 isGlyThrPheIleGlyAlaAspAspLeuAlaLys 596
Db 1748 ATGGGTATTATTGTCAAGGAAAGTGACCTTAATAAG 1782
RESULT 9
US-09-758-269-7
; Sequence 7, Application US/09758269
; Patent No. US20020104120A1
; GENERAL INFORMATION:
; APPLICANT: IUCHI, SATOSHI
; APPLICANT: KOBAYASHI, MASATOMO
; APPLICANT: SHINOZAKI, KAZUO
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
; TITLE OF INVENTION: CLEAVAGE ENZYME GENE
; FILE REFERENCE: 3914-3
; CURRENT APPLICATION NUMBER: US/09758,269
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: JP 2001-003476
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: JP 2000-010056
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1617
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1614)
US-09-758-269-7
Alignment Scores:
Pred. No.: 1,64e-97 Length: 1617
Score: 938.00 Matches: 207
Percent Similarity: 55.81% Conservative: 105
Best Local Similarity: 37.03% Mismatches: 199
Query Match: 27.98% Indels: 48
DB: 13 Gaps: 13
US-09-758-269-6 (1-599) x US-09-758-269-7 (1-1617)
Qy 64 LysGlnSerSerAsnSerProAlaIleValLysProLys---AlaLysGluSerAsn 82
Db 10 AAATCTAGTGTAGCGAGCAGCATCTCTAGTCATCTTAGCCCTCCAGGGTTTCTCC 69
Qy 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAlaGluGly 102
Db 70 TCGAAGCTTCTCGATCTTCGAGAGACTTGTTCAGGCTCATG-----TAC 114
Qy 103 PheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerValGln 122
Db 115 -----CACGATGCTTCTCCTCCTCTCCAC-----TAC 141
Qy 123 IleAlaGlyAsnPheAlaProVal---AsnGluGlnProValArgAsnLeuProVal 141
Db 142 CTCTCAGCAACTTCGCTCCATCCGTCATGAAACTCTCCCGTCAGGATCTCCCGTC 201
Qy 142 ValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnProLeu 161
Db 202 CATGATTTCTCCGAATGCTTGAATGGTGAATTTGTAGGGTTGGTCCAAACCCCAAG 261
Qy 162 HisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaValLys 181
Db 262 TTTGATGTCGTGCGTGGATATCACTGGTTTTCATGGAGATGGGATGATTCATGGGGTACGC 321
Qy 182 PheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheValGln 201
Db 322 ATCAAGATGGGAAGCTACTTATGTTTCTCGATATGTTAAGACATCAGCTCTTAAGCAG 381
Qy 202 GluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHisThr 221
```

```
Db 382 GAAGAGTTCTTCGGAGCTGCCAAATTCATGAAG---ATTGGTGACCTTAAGGGTTTTTC 438
Qy 222 GlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspProAla 241
Db 439 CGATTGCTAATGCTCAATGCTCAACAGCTGAGAACGAAGCTCAAAATATTGGACAACACT 498
Qy 242 HisGlyThrGlyValAlaAsnAlaGlyIleValTyrPheAsnGlyArgLeuLeuAlaMet 261
Db 499 TATGGAATATGGAACACTGCCAATACAGCACTCGTATATACACCATGGAAACCTTCTACATTA 558
Qy 262 SerGluAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThrVal 281
Db 559 CAGGAGGCAGATAAGCCGTACGTCACTCAAAAGTTTGGAGATGGAGACCTGCAAACTCT 618
Qy 282 GlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysValAsp 301
Db 619 GGTATAATAGATTATGACAAGAGATTGACCCACTCTCTCACTGCTCACCCAAAGTTGAC 678
Qy 302 ProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeuLys 321
Db 679 CCGGTTACGGGTGAATGTTTACATTCGGCTATTCCG---CATACGCCACCTTATCTCACA 735
Qy 322 TyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAspGln 341
Db 736 TACAGATTATCTCGAAGATGCGATTATGATGACCCAGTCCCAATTACTATATACAG 795
Qy 342 ProThrMetHisAspPheAlaIleThrGluAsnPheValValProAspGlnGln 361
Db 796 CCTATCATGATCATGATTTTCTTACTGAGACTTATGCAATCTTCATGATCTTCTCT 855
Qy 362 ValValPheLysLeuProGluMetIleArgGlyGlySerProValVal---TyrAspLys 380
Db 856 ATGCACCTTCAGGCCAAAGAAATGTTGAAGAGAGAAATATATATATCTCATTTGATCCC 915
Qy 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnLysLys 400
Db 916 ACAAAAAGGCTCGTTTGGTGTCTTCACCGCTATGCCAAGGATGAACTATGATTAGA 975
Qy 401 TrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluProGlu 420
Db 976 TGGTTTGAGCTTCCCACTGCTTTTATTTCCACACGCCCAATGCTTTGGAA-----GAA 1029
Qy 421 ThrAspGluValValIleGlySerCysMetThrProProAsp----- 435
Db 1030 GAGGATGAATGCTCTCTCATCTTGTCTGTGAGAAATCCAGATCTTCACATGTCAGT 1089
Qy 436 SerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsn 455
Db 1090 GGGAAAGTGAAAGAAAACCTGAAAATTTTGGCAACGAACGTGACGAAATGAGATTCAAC 1149
Qy 456 LeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnValAsn 475
Db 1150 ATGAAAACGGGCTCAGCTTCTCAAAAAAACTATCCGATCTGCG----- 1194
Qy 476 LeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeu 495
Db 1195 GTTGATTTTCCCAAGATCAATGAGTGCTACACCGGAAAGAAACAGACAGATACGTATATGA 1254
Qy 496 AlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAsp----- 511
Db 1255 ACAATTCTGGACAGTATCCGAAGGTATCCGGAATCATCAAGTTTGTATCTCGATGCAGAA 1314
Qy 512 -----LeuThrThrGlyGluValLysLysHisLeuTyr----- 522
Db 1315 GCTGAGACAGGAAAAGATGCTGGAAGTAGGAGTAAATATCAAGGAATATATACACCTG 1374
Qy 523 GlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGlyGlyGlyGluAsp 542
Db 1375 GGAAGAAGGCAGATATGGTTCAGAGCTATCTATGTTCCGCGTGACACAGCAGAGAAGAC 1434
Qy 543 GluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIle 562
Db 1435 GACGGTTACTTGATATCTTTGTTTCATGATGAAAAACACAGGGAATCATCGTGACTGTG 1494
```

```
Qy 563 ValAsnAlaValSerLeuGluValGlu-----AlaThrValLysLeuProSerArgVal 580
Db 1495 ATAGAGCAAAACAAATGCTCGCTGAACCGGTGGCAGTGGTGGAGCTGCCGCACAGGTC 1554
Qy 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspLeuAlaLysGlnValVal 599
Db 1555 CCATATGCTTCCCATGCTTGTGTTTATACAGAGGAACAACCTCCAGGAACAACACTCTT 1611

RESULT 10
US-09-758-269-17
; Sequence 17, Application US/09758269
; Patent No. US20020104120A1
; GENERAL INFORMATION:
; APPLICANT: IUCHI, SATOSHI
; APPLICANT: KOBAYASHI, MASATOMO
; APPLICANT: SHINOZAKI, KAZUO
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
; TITLE OF INVENTION: CLEAVAGE ENZYME GENE
; FILE REFERENCE: 3914-3
; CURRENT APPLICATION NUMBER: US/09758,269
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: JP 2001-003476
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: JP 2000-010056
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 1617
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1614)
US-09-758-269-17

Alignment Scores:
Pred. No.: 2,14e-97 Length: 1617
Score: 937.00 Matches: 207
Percent Similarity: 55.81% Conservat: 105
Best Local Similarity: 37.03% Mismatches: 199
Query Match: 29.75% Indels: 48
DB: 10 Gaps: 13

US-09-758-269-6 (1-599) x US-09-758-269-17 (1-1617)
Qy 64 LysGlnSerSerAsnSerProAlaIleValLysProLys---AlaLysGluSerAsn 82
Db 10 AAACCTCAGTGTGCGAGCATCATCTCAGTCCATCTCAGTCCCTAGACCTCCAGGGTTTCTCC 69
Qy 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAlaGluGly 102
Db 70 TCGAAGTCTTCGATCTTCGAGAGACTTCTCCTCCTCTCCAC-----TAC 141
Qy 103 PheLeuValSerHisGlyLysLeuHisProLysThrAlaAspProSerValGln 122
Db 115 -----CAGCATGCTTCTCTCCTCTCCAC----- 141
Qy 123 IleAlaGlyAsnPheAlaProVal---AsnGluGlnProValArgAsnLeuProVal 141
Db 142 CTCTCAGCAACTTCGCTCCCATCGTATGAAACTCTCCCGTCAAGGATCTCCCGCTC 201
Qy 142 ValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyValAsnProLeu 161
Db 202 CATGATTTCTTCGGAATGCTTGAATGGTGAATTTGTAGGGTTGGTCCAAACCCCAAG 261
Qy 162 HisGluProValThrGlyHisHisPhePheAspGlyAspGlyMetValHisAlaValLys 181
Db 262 TTTGATGCTGCTCGTGGATATCACTGTTGATGGAGATGGATGATTCATGGGGTAGCG 321
Qy 182 PheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheValGln 201
```



## RESULT 12

US-08-976-063C-1/c  
 ; Sequence 1, Application US/08976063C  
 ; Publication No. US20020182697A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alexander Steinbuechel; Horst Priefert; Jurgen Rabenhorst  
 ; TITLE OF INVENTION: SYNTHETIC ENZYMES FOR THE PRODUCTION OF  
 ; TITLE OF INVENTION: CONFERYL ALCOHOL, CONFERYLALDEHYDE, FERULIC ACID, VANILLIN A  
 ; TITLE OF INVENTION: ACID AND THEIR USE  
 ; NUMBER OF SEQUENCES: 42  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SPRUNG KRAMER SCHAEFER & BRISCOE  
 ; STREET: 660 White Plains Road  
 ; CITY: Tarrytown  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10591-5144  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage  
 ; COMPUTER: HP VECTRA  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: Wordperfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/976,063C  
 ; FILING DATE: 21-NOV-1997  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 196 49 655.1 (Germany)  
 ; FILING DATE: 29-NOV-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kurt G. Briscoe  
 ; REGISTRATION NUMBER: 33,141  
 ; REFERENCE/DOCKET NUMBER: Bayer 9998-CAO  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (914) 332-1700  
 ; TELEFAX: (914) 332-1844  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 32679 base pairs  
 ; TYPE: Nucleic Acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (Genomic)  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Pseudomonas sp.  
 ; STRAIN: HR199  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 3146..3997  
 ; OTHER INFORMATION: /gene= "ORF1"  
 ; US-08-976-063C-1

Alignment Scores:  
 Pred. No.: 1,12e-35 Length: 32679  
 Score: 419.00 Matches: 160  
 Percent Similarity: 37.97% Conservative: 83  
 Best Local Similarity: 25.00% Mismatches: 229  
 Query Match: 13.30% Indels: 168  
 DB: 8 Gaps: 24

US-09-758-269-6 (1-599) x US-08-976-063C-1 (1-32679)

QY 19 HisThrGlnProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSerSerLeu 38  
 |||||  
 Db 17011 CATACCGCCACCGAATAGTTGTCATTCTGAAATTTCTAAACAATAGCGCGCTCGGC 16952  
 |||||  
 QY 39 ProMetAlaSerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProPro 58  
 |||||  
 Db 16951 CCATTTTCA-----ATACCACCA 16934

QY 59 AlaLeuHisPheProLysGlnSerSerAsn-----SerProIleValVal 74  
 |||||  
 Db 16933 CCATTGCACCTCAGACACATCTTCTCCTTCAACCGCACACATGAATCCG-----GTC 16883  
 |||||  
 QY 75 LysProLysAlaLysGlnSerSerAsnThrLys-----GlnMetAsnLeu 88  
 |||||  
 Db 16882 AGTCCTAAACATAAACTCCAAACCGGCAAGCGCGCATTTAGCGCATTCAGAACACAA 16823  
 |||||  
 QY 89 PheGlnArgAlaAlaAlaAlaLeuAspAlaAlaGluGlyPheLeuValSerHisGlu 108  
 |||||  
 Db 16822 CGAGACCGTGCCATGCGGAGATTC----- 16799  
 |||||  
 QY 109 LysLeuHisProLeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAla 128  
 |||||  
 Db 16798 -----AACCGCAACGACCG-----CAATTAGTAGGAACACTTCTC 16763  
 |||||  
 QY 129 ProValAsnGluGlnProValArgArgAsnLeuProValValGlyLysLeuProAspSer 148  
 |||||  
 Db 16762 CCCACCGGTATAGAGCGAGACTTGTTCGATCTAGAGGTTGACGCGCAATCCCAATCA 16703  
 |||||  
 QY 149 IleLysGlyValTyrValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHis 168  
 |||||  
 Db 16702 ATAATGGAACTTCTACCGTATACCGCAGACCTCAAGTTACCCACAAAAATTCAC 16643  
 |||||  
 QY 169 HisPheAspGlyAspGlyMetValHisAlaValLysPheGluHisGlySerAlaSer 188  
 |||||  
 Db 16642 ACCTTCATAGATGGAGATGGAATGCCTCTGCCTTCCACTTCGAAGATGTCATGCGAC 16583  
 |||||  
 QY 189 TyrAlaCysArgPheThrGlnThrAsnArgPheValGlnGlnArgGlnLeuGlyArgPro 208  
 |||||  
 Db 16582 TTCATCAGTCGCTGGGTTAAAAACCGCTCGATTCACGCGCGAAGCACTAGCGGAAATCG 16523  
 |||||  
 QY 209 ValPheProLysAlaIleGlyLeuHisGlyHisThrGlyIleAlaArgLeuMetLeu 228  
 |||||  
 Db 16522 CTATTGGCATGTACAGAAACCCCTATACCGAGCACACAGTGTAAAGGACTA----- 16469  
 |||||  
 QY 229 PheTyrAlaArgAlaAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsn 248  
 |||||  
 Db 16468 -----GACCGCACCGTGTCCAAT 16451  
 |||||  
 QY 249 AlaGlyLeuValTyrPheAsnGlyArgLeuLeuAlaMetSerGlnAspAspLeuProTyr 268  
 |||||  
 Db 16450 ACAAGCATCATTAGCCATCAGCGCAAGTGTCTGGCGGTGAAGAAAGACAGCGCTACGCTAC 16391  
 |||||  
 QY 269 GlnValGlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAspPheAspGly 288  
 |||||  
 Db 16390 -----GAACCTGGATCCTCGT---ACACTTGAATCTCGGACACTTCGACTACGACGCG 16340  
 |||||  
 QY 289 GlnLeuGluSer---ThrMetIleAlaHisProLysValAspProGluSerGlyGluLeu 307  
 |||||  
 Db 16339 CAAGTTACGAGCCAAACCCACACCGCCCATCCAAAATATGACCCCGAAACGGGTGACTTG 16280  
 |||||  
 QY 308 PheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerPro 327  
 |||||  
 Db 16279 TTG-----TTCTTCGTTTCGCGAGCT 16259  
 |||||  
 QY 328 AspGlyThrLysSerProAspVal----- 335  
 |||||  
 Db 16258 AAGGCGGAAGCAACTCCAGACATGCGCTATTACATTGTGCAACAGCAGCGAAGGTGACA 16199  
 |||||  
 QY 336 ---GluIleGlnLeuAspGlnPro-----ThrMetMetHisAspPheAlaIleThrGlu 352  
 |||||  
 Db 16198 CATGAACCTTGGTTGAGCAGCGCTTATGGCGCATTCATGCAGCACTTTCCTCATCCCA 16139  
 |||||  
 QY 353 AsnPheValValProAspGlnGlnValValPheLysLeuProGluMetIleArgGly 372  
 |||||  
 Db 16138 AATTGGTCCATTTTCCCAATTATGCGCGCCACCAACAGCGCTG---TCCCGCTCAAGGCG 16082  
 |||||  
 QY 373 GlySerProVal---ValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLys 391  
 |||||  
 Db 16081 AAACGCAATTTATATGTGGAGCGCGAACTGGGCGAGCTACATTTGGCGGTACTC----- 16028



[illegible]

RESULT 14

```

US-09-878-574-2872
; Sequence 2872, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 2872
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(320)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3028-019-Q1-B1-B11
US-09-878-574-2872

```

## US-09-758-269-6 (1-599) x US-09-878-574-2872 (1-320)

Qy	447	SerValLeuSerGluIleAArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIle	466
Db	2	AGCGTTTAAACAGAAAGTAAGCGTGAACATGTGAAGACCGGAAAGCGAGGAGGGTGCTG	61
Qy	467	IleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeu	486
Db	62	GTG-----GAGGAAATGAACCTGGAGGCAGGATGGTGAACAGGAAAAAGGTG	109
Qy	487	GlyArgLysThrLysPheAlaTyLeuAlaLeuAlaGluProTrrProLysValSerGly	506
Db	110	GGAGAGAAAAACACGGTTTCGCATATTTGTGCATAGCGAACCGTGGCCGAAGGTGTCGGGG	169
Qy	507	PheAlaLysValAspLeuThrThrGlyGluValLysLysHisLeuTyTyrGlyAspAsnArg	526
Db	170	GTGCGAAGAGTGGACCTGGAGAGTGGGGAGTGAAGAGGCACGAGTACGGAGAGAGAAGG	229
Qy	527	TyrGlyGlyGluProLeuPheLeuPro-----GlyGluGlyGlyGluGluAspGluGly	544
Db	230	TTCGGTGGAGAGCCATTCTTCTGCCAACACGTGGAGGGAATGGGAATGAGATGAANGG	289
Qy	545	TyrIleLeuCysPheValHisAspGlu	553
Db	290	TAGTGATGGCTTTTGTGCATGACGAG	316

RESULT 15

US-09-770-696-172/c  
; Sequence 172, Application US/09770696  
; Patent No. US20010044940A1  
GENERAL INFORMATION:  
; APPLICANT: Gorlach, Jörn  
; APPLICANT: An, Yong-Qiang  
; APPLICANT: Hamilton, Carol M.  
; APPLICANT: Price, Jennifer L.  
; APPLICANT: Raines, Tracy M.  
; APPLICANT: Yu, Yang

